TRAC-M-TM-15-30
October 2015

Understanding Optimal Decision-making in War-gaming III

TRADOC Analysis Center
700 Dyer Road
Monterey, California 93943-0692

This study cost the
Department of Defense approximately
$124,000 expended by TRAC in
Fiscal Years 15.
Prepared on 20150921
TRAC Project Code # 060123.

TRAC-M-TM-15-30
October 2015

Understanding Optimal Decision-making in War-gaming III

MAJ Cardy Moten III

TRADOC Analysis Center
700 Dyer Road
Monterey, California 93943-0692

This page intentionally left blank.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)* 15-09-2015	**2. REPORT TYPE** Technical Memorandum	**3. DATES COVERED** *(From - To)*
4. TITLE AND SUBTITLE Understanding Optimal Decision-making in War-gaming III		**5a. CONTRACT NUMBER**
		5b. GRANT NUMBER
		5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S) MAJ Cardy Moten III		**5d. PROJECT NUMBER** 060123
		5e. TASK NUMBER
		5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) TRADOC Research Analysis Center, Monterey, CA Naval Postgraduate School, Monterey, CA		**8. PERFORMING ORGANIZATION REPORT NUMBER** TRAC-M-TM-15-030
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Army Research Office, Research Triangle Park, NC		**10. SPONSOR/MONITOR’S ACRONYM(S)**
		11. SPONSOR/MONITOR’S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The purpose of this memorandum is to provide documentation of research for the Army Research Office (ARO) by the TRADOC Analysis Center, Monterey (TRAC-MTRY). The focus of the research is to develop a model that represents the relationship between neurophysiological metrics and optimal decision making. The research team modified two well-known psychological tests for a military context. The Iowa Gambling Task (IGT) was modified to assess reinforcement learning and the Wisconsin Card Sorting Test (WCST) was modified to assess cognitive flexibility. The tests were administered to 34 military officers across all services. Based on the results of the IGT and WCST, the research team also developed the Cognitive Alignment With Performance Targeted Training Intervention Model (CAPTTIM) to assess the relationship between a subject's cognitive state and their observed performance. Through analyzing reinforcement learning and cognitive flexibility, the CAPTTIM can be used to provide a real-time notification of when a training intervention is required and the type of training intervention necessary. The results indicate that the modified versions of the IGT and WCST along with the CAPTTIM can be as an objective assessment tool in conjunction with other virtual and live military decision making training.

15. SUBJECT TERMS

Training intervention, cognitive state, regret, neurophysiological measures, Wisconsin Card Sorting Test, Iowa Gambling Task

16. SECURITY CLASSIFICATION OF: Unclassified			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON** MAJ Cardy Moten III
a. REPORT Unclassified	**b. ABSTRACT** Unclassified	**c. THIS PAGE** Unclassified	UL	115	**19b. TELEPHONE NUMBER** *(include area code)* 831-656-2452

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

This page intentionally left blank.

Table of Contents

This page intentionally left blank.

1. **Purpose.** The purpose of this memorandum is to provide documentation of research for the Army Research Office (ARO) by the TRADOC Analysis Center, Monterey (TRAC-MTRY). The focus of this phase is to document research on the development of a model that represents the relationship between neurophysiological metrics and optimal decision making.

2. **Background.** The U.S. Army published its operating concept in October of 2014. The purpose of this concept is to describe how the Army will operate at the strategic, operational, and tactical level without knowing much about the future environment, location, and enemy.[1]In order to accomplish this objective, the training for Army officers has to focus on adaptive decision making through realistic training in actual and virtual environments.[2] Currently, the metrics used in training to evaluate the decision making of officers is subjective, and little is known about how military officers make optimal decisions. A potential solution to this problem is to combine human-in-the-loop wargames with behavioral and neurophysiological measures.

3. **Methodology.**The research team modified two well-known psychological tests for a military context. The Iowa Gambling Task (IGT) was modified to assess reinforcement learning.[3] The Wisconsin Card Sorting Test (WCST) was modified to assess cognitive flexibility.[4] The tests were administered to 34 military officers across all services. Kennedy et al. discuss in detail the modification of these tests and the results of their research.[5] Based on the results of the IGT and WCST, the research team also developed the Cognitive Alignment With Performance Targeted Training Intervention Model (CAPTTIM) to assess the relationship between a subject's cognitive state and their observed performance. Through analyzing reinforcement learning and cognitive flexibility, the CAPTTIM can be used to provide a real-time notification of when a training intervention is required and the type of training intervention necessary.[6] This is done through using quantitative statistical methods to determine if a decision maker is in an exploration versus exploitation cognitive state and if they are yielding the optimal decision performance while in that particular state. In this research that decision performance metric is the amount of regret, which we define as the difference between the maximum benefit that could be received at a particular state minus the value of the benefit actually obtained. An exploration cognitive state indicates the subject is

[1]U.S., Department of the Army Training and Doctrine Command. *TRADOC Pamphlet 525-3-1, The U.S. Army Operating Concept: Win In a Complex World.* Washington DC: Government Printing Office, October 2014.

[2]Ibid.

[3]Antoine Bechara et al. "Insensitivity to future consequences following damage to human prefrontal cortex". In: *Cognition* 50.1 (1994), pp. 7–15.

[4]David A Grant and Esta Berg. "A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem." In: *Journal of experimental psychology* 38.4 (1948), p. 404.

[5]Quinn Kennedy, Peter Nesbitt, and Jon Alt. "Assessment of Cognitive Components of Decision Making with Military Versions of the IGT and WCST". in: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting.* Vol. 58. 1. SAGE Publications. 2014, pp. 300–304.

[6]Quinn Kennedy et al. *Cognitive Alignment with Performance Targeted Training Intervention Model: CAPTTIM.* tech. rep. Monterey, California. Naval Postgraduate School, 2015.

more of a naïve decision maker and needs more information on their environment.[7] An exploitation cognitive state indicates the subject is more experienced and has figured out the optimal alternative and does not consider any other sub-optimal alternative from that point on.[8]

4. **Progress.** The following is a summary of the documentation produced by the research team during FY 15. Nesbitt et al. submitted and received acceptance of their manuscript to the Journal of Military Psychology.[9] Kennedy et al. presented their findings on the how their modified version of the IGT, called the convoy task, can be used to help screen for traumatic brain injury at the 2015 meeting of the International Nueropsychological Society[10] (See Appendix B). Critz utilized change point analysis to develop and validate "a threshold that delineated near-optimal and suboptimal decision performance with the metric, regret, and categorize the combination of cognitive state and decision performance into CAPTTIM on a trial-by-trial basis"[11] (See Appendix C). Moten et al. prepared a manuscript for publication on their analysis of the WCST[12] (See Appendix D).

5. **Results.** Nesbitt et al. found that their modified version of the IGT tested the same cognitive function as the original task and that using regret as an additional metric provided a suitable assessment of low and high performers.[13] Kennedy et al. concluded that poor performance on the convoy task and low latency measures are a potential indicator of traumatic brain injury. They also supported these findings with the use of EEG and eye-tracking equipment.[14] Critz determined that using regret along with latency can help a military trainer determine the training deficiencies of a military decision maker in a simple wargame.[15] Moten et al. concluded that all participants who achieved a shift in sorting rule demonstrated adequate cognitive flexibility. However, participants that did not complete all the required sorting rules changed their sorting strategy too soon within a series, resulting in a high quantity of random errors.[16]

6. **Future Work** Kennedy et al. are scheduled to present their findings on their development of CAPTTIM at the 2015 Human Factors Ergonomics Society annual meeting[17]

---

[7] Kennedy, Nesbitt, and Alt, "Assessment of Cognitive Components of Decision Making with Military Versions of the IGT and WCST", op. cit.

[8] Ibid.

[9] Peter Nesbitt et al. "Iowa Gambling Task modified for military domain." In: *Military Psychology* 27.4 (2015), p. 252.

[10] Quinn Kennedy et al. "Can a simple wargame provide an unobtrusive indicator of TBI?". In: *43rd Annual Meeting of the International Neuropsychological Society (INS)*. 2015.

[11] John Critz. "Understanding Optimal Decision Making". MA thesis. Naval Postgraduate School, 2015.

[12] Cardy Moten III et al. "Analysis of Performance on a Modified Wisconsin Card Sorting Test for the Military".

[13] Nesbitt et al., "Iowa Gambling Task modified for military domain.", op. cit.

[14] Kennedy et al., "Can a simple wargame provide an unobtrusive indicator of TBI?", op. cit.

[15] Critz, "Understanding Optimal Decision Making", op. cit.

[16] Moten III et al., "Analysis of Performance on a Modified Wisconsin Card Sorting Test for the Military", op. cit.

[17] Quinn Kennedy et al. "Cognitive alignment with performance targeted training intervention model: CAPTTIM". in: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. 2015.

and are preparing a manuscript regarding the validation of the CAPTTIM model with neuropsychological data. Moten et al. will present their insights into the modification of the WCST at the 2015 Informs conference and they are preparing a manuscript detailing the nonparametric techniques used in their analysis. Marine Major Travis Carlson will prepare a masters thesis on the development of a real-time training intervention model based on CAPTTIM.

This page intentionally left blank.

Appendix A
Study Plan

Problem Statement

To investigate the role between neurophysiological indicators and optimal decision-making in the context of military decision making scenarios as represented in human-in-the-loop wargaming simulation experiments.

Project Team

Sponsor Agency:	Dr. Virginia Pasour Biomathematics Program U.S. Army Research Office, Research Triangle Park, NC virginia.b.pasour.civ@mail.mil
TRAC Lead:	Cardy Moten III MAJ, LG/FA49 TRADOC Analysis Center - Monterey cardy.moten3.mil@mail.mil
Primary Investigator:	Dr. Quinn Kennedy Operations Research Department Naval Postgraduate School, Monterey, CA mqkenned@nps.edu
NPS Faculty:	LTC Jonathan Alt Operations Research Department Naval Postgraduate School, Monterey, CA jkalt@nps.edu

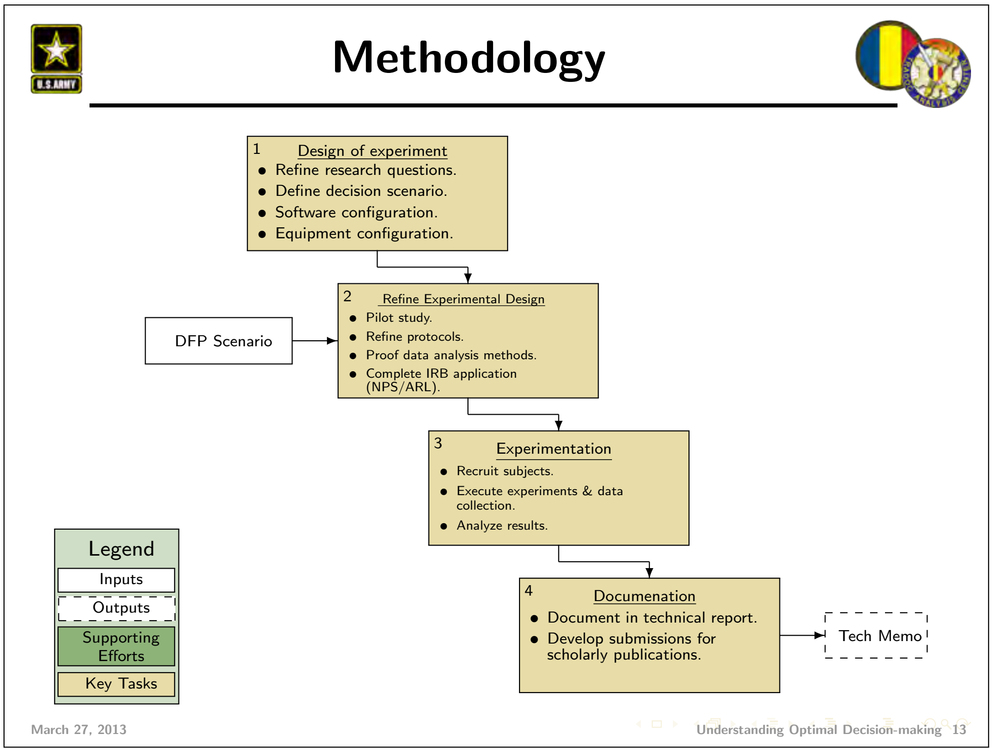
Constraints, Limitations, & Assumptions

- **Constraints**
  - The total budget for this phase of the project is $96K.
  - Phase III must be complete no later than 30 December 2015.
- **Limitations**

  - Will limit initial experimentation to discrete decision situations or with limited exposure of sequential tasks.
  - Subjects limited to those officer students available at NPS.

- **Assumptions**
  - Results of experimentation with available subject pool will be sufficient to provide insight into study issues.

Methodology



Timeline

APR 14	Submit IGT and WCST modification paper to the Human Factors and Ergonomics Society (HFES)
OCT 14	ODM II IPR
OCT 14	Present findings at the HFES annual meeting.
DEC 14	CAPTIIM Tech Report complete.

Appendix B
Conference Poster

The following page displays the conference poster presented by Kennedy et al. at the 43rd Annual Meeting of the International Neuropsychological Society.

Can a simple wargame provide an unobtrusive indicator of TBI? A case study

Quinn Kennedy, PhD, Maheen Adamson, PhD, Jesse Huston & Major Peter Nesbitt
Naval Postgraduate School and VA Palo Alto Healthcare System/Stanford University

ABSTRACT

Objective: To test the efficiency of a wargame to indicate TBI among active duty officers. **subjects and Methods:** A military version of the Iowa Gambling task, the convoy task, was created to measure reinforcement learning and military decision making. The goal of the convoy task was to maximize total damage score by minimizing friendly damage and maximizing enemy damage incurred. Over 200 trials, 34 officers (aged 29 - 45 yrs) tried to learn the best route to send convoys. After each decision, officers received immediate feedback regarding friendly, enemy, and total damage. In addition to latency response (LR), attention to feedback also was measured by eyetracking. LR was defined as the proportion of trials in which officer's decision time immediately after receiving feedback of moderate to heavy friendly damage was greater than 2 sd above their baseline time. Seventeen officers also completed a self-report Traumatic Brain Injury (TBI) survey. **Results:** One officer (age = 45; deployment time= 28 months) had moderate TBI and was used for this case study. The decision performance, response to negative feedback, and eye scan pattern in the officer with TBI differed dramatically from the overall sample. First, his total damage score was 650 (sample mean = 2456.1 (sd = 1724.0). Second, in the sample, LR was positively associated with total damage score (r = 38, p = .03). The LR rate for the officer with TBI was 0% (sample mean = 18.2%). Third, preliminary eyetracking results in the sample revealed that looking at enemy damage was not correlated with total damage score; this officer spent almost double the amount of time looking at enemy damage (12% vs. mean = 6.4%). Note that the officer with TBI scored in the 90% on Trails A and B, and within normal range for digit span forwards and backwards. **Conclusions :** This case study illustrates the utility and efficiency of a military-specific task that can be used to indicate TBI in a population of active duty officers.

METHOD

subjects: Nine Army, 11 Marine Corps, 10 Navy, 3 Coast Guard, 1 Air Force officers (n = 34), with mean age of 35.1 (4.9) years participated in the study. They had a mean time in service of 12.7 (4.4) years, and mean time deployed of 19.6 (12.1) months, with a mean of 38.0 (25.2) months since their last deployment. Most subjects were male (88.3%) and had 20/20 vision (85.3%). All subjects had college degrees and were in pursuit of Master's degrees.

Main Measures

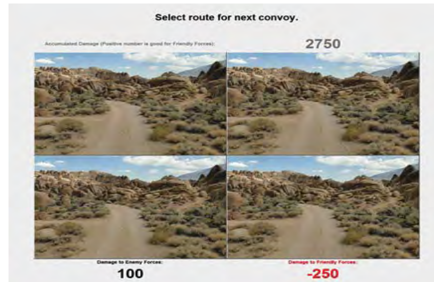
Convoy task: In a military version of the Iowa Gambling Task (Bechara et al 1994), officers view four identical routes and must decide the route on which to send their convoys (see Figure 1). Depending on their decision, officers can inflict enemy damage (good) and sometimes receive friendly damage (bad). The objective is to maximize the total damage score by maximizing enemy damage and minimizing friendly damage. Routes 3 and 4 provide the best long-term total damage scores (see Table 1). Every 10th time that Route 2 is selected, heavy friendly damage occurs. All subjects start with 2000 damage. Main performance measures are:

- Total damage score: Enemy damage minus friendly damage.
- Advantageous selection bias: Proportion of good routes selected minus the proportion of bad routes selected.
- Latency response: Proportion of trials in which officer's decision time immediately after receiving feedback of moderate to heavy friendly damage was greater than 2 sd above their baseline time.

Table 1 Summary statistics for the damage that can occur for each route during the convoy task. Negative numbers indicate friendly damage; positive numbers indicate enemy damage.

	Route 1	Route 2	Route 3	Route 4
Minimum	-250	-1250	0	-200
25%	-150	100	0	50
Median	25	100	25	50
Mean	-25	-25	25	25
75%	100	100	50	50
Maximum	100	100	50	50

Figure 1 The convoy task. Through reinforcement learning, military personnel determine which route has the best combination of damage to enemy forces (good) and damage to friendly forces (bad). In this example, the subject selected Route 1.



Procedures: After providing informed consent, officers completed 200 trials of the convoy task while their eye gaze and brain activity were monitored via eyetracking and EEG technology. They also completed digit span forwards and backwards, Trails A and B, and a post-task survey. Officers were later contacted and asked to complete the Ohio State University TBI Identification Method Short form survey. Seventeen subjects completed the TBI survey.

RESULTS

One subject's responses to the TBI survey were consistent with moderate TBI. His performance on the convoy task differed dramatically from the overall sample, despite having high scores on the cognitive tests (see Table 2 and Figure 2). On the post-task survey, he reported that Route 2 was the safest route and showed almost no behavioral response to receiving heavy friendly damage that occurs on Route 2 (see Figure 3). He also reported Route 4 as the 2nd most dangerous route.

Table 2 Comparison between the overall sample and the TBI subject on convoy task, eyetracking, and cognitive measures. Bolded numbers: TBI performance is outside the 95% CI.

	Whole Sample excluding TBI subject (95% CI)	TBI subject
Performance variables		
Total damage score	2456.1 (1867.9 – 3044.3)	**650**
Friendly damage trials,#	51.4 (47.6 – 55.2)	**42**
Heavy friendly damage trials, #	6.6 (5.7 – 7.5)	**10**
Advantageous selection bias	9.3 (-12.1 – 30.69)	**-38**
Latency Response	18.8 (14.6 – 22.8)	**0**
Eyetracking: Percent dwell time per region		
Total damage, %	5.6 (-2.6 – 13.8)	2.7
Friendly damage, %	16.7 (3.4 – 30.1)	18.7
Enemy damage, %	6.4 (-2.4 – 15.2)	11.9
Routes, %	71.4 (55.2 – 87.6)	66.7
Cognitive measures		
Trails A (age normed), %	56.0 (39.1 – 72.9)	**90**
Trails B (age normed), %	70.9 (55.4 – 86.4)	**90**
Digit Forwards	11.4 (10.7 – 12.1)	**13**
Digit Backwards	9.4 (8.6 – 10.2)	**15**

Figure 2 Comparison of total damage per trial for the TBI subject (green line) compared to mean total damage across all subjects per trial (blue line) with 95% confidence interval (red dashed lines).

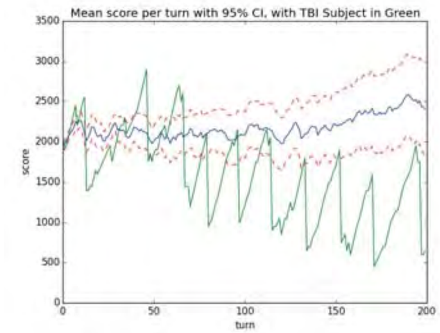


Figure 3. Latency response by trial for the TBI subject, color coded by the amount of friendly damage on the previous trial.

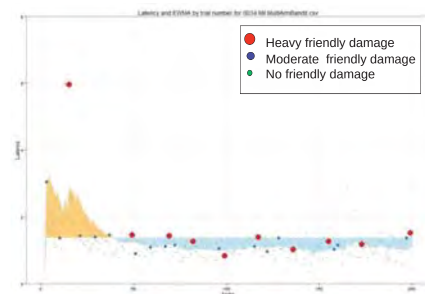
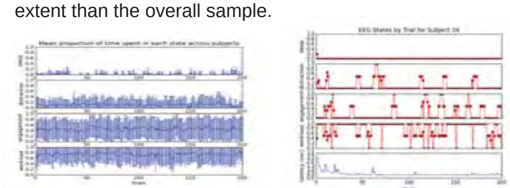


Figure 4 EEG results suggest that the TBI subject experienced high cognitive workload to a greater extent and was engaged to a lesser extent than the overall sample.



Summary and Discussion

Results from this case study suggest:

- Poor performance on the convoy task may be a better indicator of TBI status than common cognitive measures.
- The use of latency response, a simple behavioral response measure to bad outcomes, may also indicate TBI.
- EEG measures may detect unusual patterns of cognitive state experienced during the task.
- Eyetracking measures can detect poor attention allocation.

This material is based upon work supported in part by the Army Research Office (62626-NS). The content is solely the responsibility of the authors and does not necessarily represent the official views of the Army Research Office.

Appendix C
Understanding Optimal Decision-Making Thesis

The following pages consists of the thesis by U.S. Marine Captain John Critz. Distribution in unlimited.

NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

UNDERSTANDING OPTIMAL DECISION-MAKING

by

John W. Critz

June 2015

Thesis Advisor:	Quinn Kennedy
Co-Advisor:	Jon Alt

This thesis was performed at the MOVES Institute

Approved for public release; distribution is unlimited

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE	*Form Approved OMB No. 0704–0188*

1. AGENCY USE ONLY *(Leave blank)*	**2. REPORT DATE** June 2015	**3. REPORT TYPE AND DATES COVERED** Master's Thesis

4. TITLE AND SUBTITLE
UNDERSTANDING OPTIMAL DECISION-MAKING

5. FUNDING NUMBERS

6. AUTHOR(S) John W. Critz

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Postgraduate School
Monterey, CA 93943–5000

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES)
N/A

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. IRB Protocol number ____N/A____.

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (maximum 200 words)

The military has realized that their most valuable and adaptable assets are its leaders. Understanding optimal decision-making will allow the military to more effectively train its leaders. The Cognitive Alignment with Performance Targeted Training Intervention Model (CAPTTIM) was developed to aid the training of optimal decision making. CAPTTIM determines when decision performance (categorized as near-optimal or suboptimal) is aligned or misaligned with cognitive state (categorized as exploration or exploitation): when someone thinks they have figured out the task (exploitation cognitive state), is their decision performance actually near optimal? Prior research categorized subjects' cognitive states as exploration or exploitation, but the delineation of decision performance had yet been done. The primary focus of this thesis was to use pre-collected and de-identified data to (1) determine and validate a threshold that delineated near-optimal and suboptimal decision performance with the metric, regret, and (2) categorize the combination of cognitive state and decision performance into CAPTTIM on a trial-by-trial basis. A change point analysis of regret provided an effective threshold delineation of decision performance across all subjects. Visualization techniques were employed to categorize decision and cognitive state data into CAPTTIM on a trial-by-trial basis. Thus, CAPTTIM was validated as a means of understanding decision-making.

14. SUBJECT TERMS
optimal decision-making, regret, Iowa gambling task, exponentially weighted moving average, change point analysis

15. NUMBER OF PAGES
97

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	**18. SECURITY CLASSIFICATION OF THIS PAGE**	**19. SECURITY CLASSIFICATION OF ABSTRACT**	**20. LIMITATION OF ABSTRACT**
Unclassified	Unclassified	Unclassified	UU

NSN 7540–01–280–5500

Standard Form 298 (Rev. 2–89)
Prescribed by ANSI Std. 239–18

THIS PAGE INTENTIONALLY LEFT BLANK

UNDERSTANDING OPTIMAL DECISION-MAKING

John W. Critz
Captain, United States Marine Corps
B.S., University of North Carolina at Charlotte, 2008

Submitted in partial fulfillment of the
requirements for the degree of

**MASTER OF SCIENCE IN
MODELING, VIRTUAL ENVIRONMENTS, AND SIMULATION (MOVES)**

from the

**NAVAL POSTGRADUATE SCHOOL
June 2015**

Author: John W. Critz

Approved by: Dr. Quinn Kennedy
Thesis Advisor

LTC Jon Alt
Co-Advisor

Chris Darken
Chair, MOVES Academic Committee

Dr. Peter Denning
Chair, Department of Computer Science

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

The military has realized that its most valuable and adaptable assets are its leaders. Understanding optimal decision-making will allow the military to more effectively train its leaders. The Cognitive Alignment with Performance Targeted Training Intervention Model (CAPTTIM) was developed to aid the training of optimal decision making. CAPTTIM determines when decision performance (categorized as near-optimal or suboptimal) is aligned or misaligned with cognitive state (categorized as exploration or exploitation): when someone thinks they have figured out the task (exploitation cognitive state), is their decision performance actually near optimal? Prior research categorized subjects' cognitive states as exploration or exploitation, but the delineation of decision performance had yet been done. The primary focus of this thesis was to use pre-collected and de-identified data to (1) determine and validate a threshold that delineated near-optimal and suboptimal decision performance with the metric, regret, and (2) categorize the combination of cognitive state and decision performance into CAPTTIM on a trial-by-trial basis. A change point analysis of regret provided an effective threshold delineation of decision performance across all subjects. Visualization techniques were employed to categorize decision and cognitive state data into CAPTTIM on a trial-by-trial basis. Thus, CAPTTIM was validated as a means of understanding decision-making.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

First and foremost I would like to thank God, my wife Sarah, and my daughters, Eva and Elena, for being by my side and supporting me through the completion of my Master's of Science in Modeling of Virtual Environments and Simulations.

I also owe a huge debt of gratitude to Dr. Quinn Kennedy and LTC Jon Alt for their unwavering support and assistance. Without their keen insight and knowledge this thesis would not have been possible. Thank you for seeing me through to the finish.

I am forever in debt of MAJ Moten Cardy for his invaluable assistance in the scripting and writing of the R Studio code that performed the analysis portion of this thesis. Without his innovation and ability to instruct, I would have been totally lost.

A special thank you is also in order for all the Modeling and Simulations professors for their instrumental instruction and tireless preparation that made this thesis possible. Your dedication to educating military members is unmatched and much appreciated.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

A. BACKGROUND

Understanding optimal decision-making is an extremely complex task, but one that the military is currently trying to accomplish. The focus on decision-making is being renewed in an effort to not only understand the processes involved in decision-making, but also improve decision-making among service members. The goal of improving effective decision-making is to increase the combat effectiveness of the military. The last 14 years of combat operations in Afghanistan and Iraq have illustrated the necessity for military leaders to be adaptable, agile, and able to operate in a threat environment that spans irregular and regular warfare, terrorist activity, and at times even governance (Lopez, 2011). The combat environment has always been complex; however, in a non-conventional environment (irregular warfare), that complexity is increased exponentially. The recent and ongoing conflicts in Iraq and Afghanistan illustrate the importance of developing leaders with the cognitive flexibility to learn from feedback from their environment to improve decision performance. in these two conflicts leaders sometimes drew false conclusions about the effectiveness of their operations by attending to historically used measures of performance, such as enemy attrition. From personal experience, a lot of confusion occurred when high enemy body counts were not associated with victory or decreased violence. There was an inability to recognize through trial and error and reinforcement learning that the current approach was not successful. A lot of reinforcement of failure occurred, because of this lack of understanding. Had the military understood optimal decision-making better, this reinforcement of failure could have possibly been avoided by making the decision maker more adaptable, agile, and aware of the complex nuances of the counter-insurgency environment.

The military is in an ideal position to evaluate decision-making among current service members who have spent the last eleven years engaged in combat operations in Iraq and Afghanistan. With this wealth of combat

knowledge contained within current active duty service members, the military can glean decision-making patterns from experienced decision makers. These patterns can then be analyzed in order to better understand how experienced decision makers arrive at optimal or near-optimal decisions. Once this process is understood, then the military can (1) improve combat effectiveness by developing programs to improve decision making among its current leaders and (2) instruct future leaders on optimal decision making to improve their leadership potential.

The primary goal of understanding optimal decision-making is to develop training aids to instruct naïve service members in an effort to shorten the experiential knowledge required to develop effective decision-making practices in combat. Another goal of these training aids is to provide the instructor with insight into the trainee's decision-making process. Such training aids would benefit instructor to trainee interaction and provide insight on timing and type of intervention required by the instructor.

Kennedy, Nesbitt, and Alt (2014) developed a training intervention model called Cognitive Alignment with Performance Targeted Training Intervention (CAPTTIM). This model seeks to determine if a trainee's cognitive state is aligned or misaligned with their actual performance. The model utilizes latency in decision-making to determine the trainee's cognitive state; however, no "generic" metric for determining actual performance has been researched. This thesis seeks to determine an appropriate threshold that delineates between high and low regret. Determining a threshold between high and low regret is an essential step before the model can be tested.

B. REINFORCEMENT LEARNING IS NECESSARY TO REACH OPTIMAL DECISION-MAKING

One cognitive characteristic necessary for military personnel to reach optimal decision-making is reinforcement learning, the ability to learn from trial and error (Sutton & Barto, 1998). Reinforcement learning is necessary when there is a high degree of uncertainty. High levels of uncertainty are associated

with combat operations and environments, in which limited intelligence is known about the situation, but high stake decisions still have to be made. In these situations the military leader makes a “best guess” decision based on experience and training. Current reinforcement learning tests, which are typically computerized laboratory tests, do not completely capture the stressors, uncertainty, and high risk conditions of decisions made in combat (Nesbitt, Kennedy, & Alt, 2015). For example, the Iowa Gambling Task (IGT) (Bechara, Damasio, Damasio, & Anderson,1994), a very common test of reinforcement learning that has been used in hundreds of psychology studies (Krain, Wilson, Arbuckle, & Castellanos 2006), entails selecting cards from four different decks in a low stress, low stakes, game playing environment. This shortfall has led to the need to create realistic military scenarios and simple wargames that elicit reinforcement learning (Nesbitt et al., 2013). Therefore, Kennedy et al (2014) modified the IGT to mirror a military environment.

1. The Iowa Gambling Task

The IGT is a well-known psychology task that elicits reinforcement learning (Bechara et al., 1994) and has been used in hundreds of studies (Krain et al., 2006). Subjects are given a loan of $2,000, presented four decks of cards (decks A-D) face down, and asked to make selections that result in maximizing profit. Figure 1 shows a screen shot of the IGT setup.

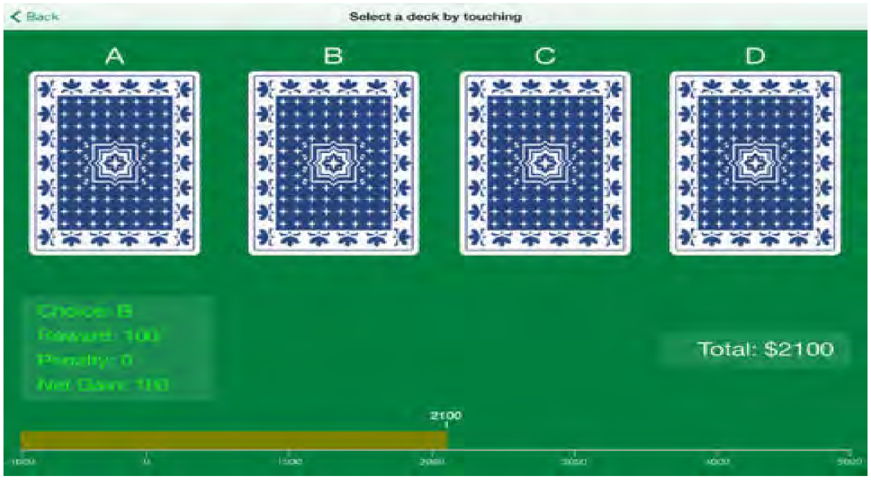


Figure 1. The Iowa Gambling Task screenshot (from Sacchi, 2014).

Each deck has a scheduled dollar payout and penalties that the subject receives depending on their deck selection. The payout amount as well as the severity and frequency of the penalty, differs from deck to deck. Subjects can change the order of their selection at any time and can choose solely from a single deck if they so desire. Through reinforcement learning, healthy subjects eventually discover that decks A and B result in long term losses, despite having higher initial payouts (Bechara et al., 1994). They then realize that, despite lower initial payouts, decks C and D result in long-term gains. Performance is measured by total money won and advantageous selection bias. Advantageous selection bias is calculated by subtracting the number of poor decisions (decks A and B) from the number of good decisions (decks C and D).

Appendix A lists the payout schedule for each deck over the 100 trials. It is important to note that the payout schedule does not reset after each card selection. Until a subject selects a particular deck, the payout for that deck remains the same. For example, Deck B has a negative 1250 penalty every tenth turn but the highest payouts otherwise; the subject cannot game the system by choosing Deck B nine times, but a different deck on the tenth turn, return to Deck B on the 11th turn in an attempt to avoid the negative 1250 penalty.

2. Convoy Task

The IGT was modified into the convoy task to reflect the risks and scenarios faced in a military environment, while mirroring the reinforcement learning elicited by the IGT. In the convoy task each subject selects a route on which to send a convoy and is given a choice between four different convoy routes. The task entails 200 trials of these decisions. At the end of each trial the subject is given immediate feedback with three separate pieces of information: a reward, a penalty, and a running total (Nesbitt et al., 2013). The reward is called *Damage to Enemy Forces*, the penalty is called *Damage to Friendly Forces,* and the running total is called *Total Damage* (Nesbitt et al., 2013). *Damage to Friendly Forces* is analogous to a loss of money in the IGT, while *Damage to Enemy Forces* is analogous to a gain of money. *Total Damage* is analogous to the loan amount and winnings in the IGT. The convoy route selection task's feedback values were adopted from the original IGT payout schedule (see Appendix A). Subjects are instructed that their goal is to maximize the total damage score by minimizing friendly damage and maximizing enemy damage. Like the IGT, subjects should learn through reinforcement learning that routes one and two are bad and routes three and four are good. Data collected from the 34 subjects who participated in the convoy task confirmed that it elicits reinforcement learning (Kennedy et al., 2014).

3. Cognitive Alignment with Performance Targeted Training Intervention

In analyzing data from the 34 subjects that participated in the convoy route task, Kennedy et al. (2015) developed a training intervention model called Cognitive Alignment with Performance Targeted Training Intervention (CAPTTIM) (see Figure 2). This model determines whether a person's cognitive state is aligned or misaligned with actual performance. The model delineates two cognitive states, exploration and exploitation. Exploration is defined as naïve decision-making, in which a person is seeking to further their understanding of the environment by gathering information. Exploitation is defined as experienced

decision-making, in which a person believes that they have attained enough information to begin acting upon that knowledge. The model quantitatively characterizes exploration and exploitation by variability in latency times on making each decision (Fricker, 2010). A standard deviation for each subject was calculated utilizing only the latency times on their decisions that resulted in no damage. Variability greater than twice the subject's standard deviation is considered exploration, whereas variability less than twice the standard deviation is considered exploitation. However, changes in latency time variability provided no measure of actual performance for the individual.

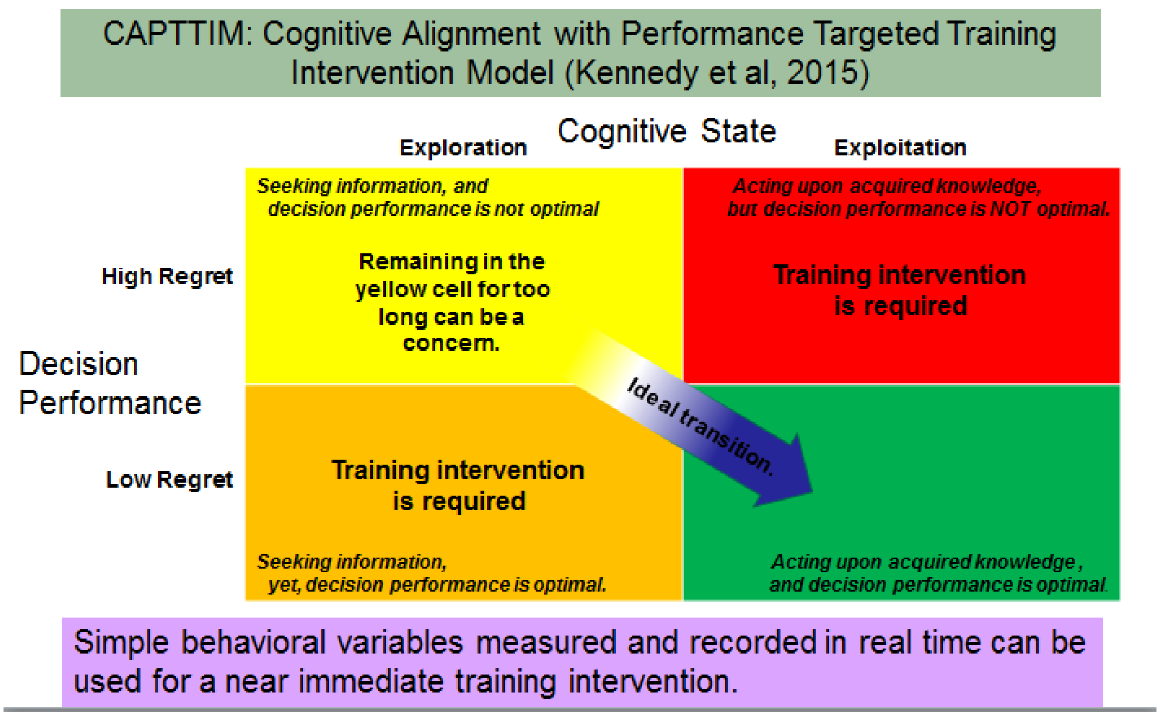


Figure 2. The combination of cognitive state and actual decision performance indicates whether a trainee's cognitive state is aligned or misaligned with actual performance. When misalignment occurs, it indicates the need for a training intervention (from Kennedy, 2015).

Actual performance is measured by regret. Regret is quantified as the difference between the maximum possible payout for a particular trial, and the actual received payout for a particular trial (Agrawal, 1995). Because the payout

schedule is consistent from individual to individual, their deviation from the optimum path can be measured. However, a threshold delineating high from low regret has not been calculated yet.

The convoy route task has a specific sequence of payouts, providing the ability to know at any point in the sequence of trials which route provides the most advantageous reward (Nesbitt et al., 2015). Because the best reward is known, it is possible to calculate the difference between the best reward and the subject's received reward at that specific trial in the convoy route selection task. This difference is defined as regret.

Regret is an absolute performance metric that provides the ability to compare actual performance of the subject with their cognitive state. If the subject's performance is misaligned with their cognitive state then the instructor can intervene and make the appropriate correction. This is very similar to Type I and Type II error from statistics. The subject's performance can be correctly aligned with their cognitive state, which is the ideal transition that is captured in CAPTTIM. Otherwise the subject is making incorrect exploitation decisions believing them to be correct (false positive), or they are making the correct decision, but do not know that they are making the correct decision (false negative). Either of the latter two options requires instructor intervention. The possibility of being able to align a trainee's cognitive state with actual performance is consistent with what the military is trying to accomplish in their pursuit of understanding optimal decision-making.

C. REGRET

Regret is used in numerous fields ranging from computer science, machine learning, and even the medical field. It is very easily applied to scenarios, like the IGT, where the optimum decision is known. For the medical field it is applied retrospectively to describe the diagnosis or misdiagnosis of patients (Djulbegovic, Elqayam, Reljic, Hozo, Miladinovic, Tsalatsanis, Kumar, Beckstead, Taylor, & Cannon-Bowers, 2014). An interesting application from this

publication that directly relates to the research question of this thesis is how much regret affects future decisions (Djulbegovic et al., 2014).

The defining principle of regret is that if you minimize regret, then you are converging on the correct decision, or for multi-arm bandit scenarios, the correct slot machine (Agrawal, 1995). This principle will be directly applied to this thesis to determine a subject's performance and determine if their performance is aligned or misaligned with their cognitive state. In layman's terms, is the subject making the right decision ignorantly, making the wrong decision thinking it is the correct decision, or do they transition correctly?

Most utilization of the principle of regret has been on analyzing its impact on decision-making or convergence on a decision in a multi arm bandit scenario. No articles could be found that discussed using regret as a method of measuring performance in the way that it is being proposed in this thesis. Other papers use regret as an additional factor in an expected utility function in an attempt to explain behaviors and choices (Bell, 1982).

Bell gives an illustrative anecdotal example of regret. He describes a farmer who has a field of crops that are not yet ready to be harvested. A buyer approaches the farmer and offers him five dollars a bushel for his produce. The farmer knows that, depending on the harvest, his produce could sell for as much as seven dollars a bushel or as little as three dollars a bushel. The farmer is faced with two potential forms of regret: (1) where he accepts the five-dollar-a-bushel offer and the harvest yields a seven-dollar-a-bushel product, (2) he refuses the five-dollar-a-bushel offer and the harvest yields a three-dollar-a-bushel product. Bell then describes how these two forms of regret have very different effects on differing subjects. For some subjects, the fear of losing two dollars per bushel, in the event of an inferior crop, influences their decision much more than the possibility of gaining an extra two dollars per bushel (Bell, 1982). Bell then highlights this phenomenon later on in his paper, when he discusses the utility function. In this example, he discusses how a person might "feel" greater regret between an outcome of $1,000 and $2,000 than an outcome of

$1,000,000 and $1,001,000, despite the fact that both gained or lost $1,000 (Bell, 1982). He discusses how the increment is not "felt" the same between both outcomes (Bell, 1982). Bell (1982) additionally made the following comment that is applicable to this thesis and could possibly explain decisions made by subjects: "At an extreme, a decision maker who has severe problems with regret may sometimes prefer to have only a single alternative offered than a choice among two or more" (p. 969). This idea could possibly explain certain subjects' behavior and their decision to only select certain routes, rather than exploring all options.

Bell additionally looked at regret to explain behaviors and gives anecdotal examples in the realm of insurance and gambling. "The consequence with the largest regret is that in which you choose not to bet, but hear that you would have won" (Bell, 1982, p. 971). If an individual decides not to bet on the horse with long odds, he or she experiences a high amount of regret if that horse wins (Bell, 1982). If you bet on the same lottery number for an extended period of time, the thought of that being the winning number as soon as you stop choosing it could be strong enough to encourage you to continue gambling (Bell, 1982). Bell argues that regret can be used to justify risk-prone behavior (gambling) and risk-averse behavior (purchasing insurance) on the part of the same decision maker (Bell, 1982). For risk-averse behavior, subjects are willing to accept the regret associated with paying for insurance, but never making a claim (Bell, 1982).

Regret is an effective performance metric in tasks in which the payout or reward is known for each decision. For this reason, it is a common performance metric used in gambling scenarios, specifically with multi-arm bandit gambling scenarios (Nesbitt et al., 2015). In these scenarios, the optimum path can be determined. Deviations from this optimum path can be quantified by this notion of regret. We now provide an example of how regret is calculated in a scenario in which the optimum path can be determined—the convoy task payout schedule (Figure 3). In this excerpt, if a subject chooses Route 4 on trial 1, their regret will be 100 – 50 = 50, because the optimum choice was either Route 1 or Route 2.

If the subject chooses Route 4 again on trial 2, their regret will be 100 – (-250) = 350, because the optimum choice was still either Route 1 or Route 2. If the subject chooses Route 2 on trial 3, their regret will be 100 – 100 = 0, because Route 2 was one of the optimum choices. If by trial 9 all routes have been selected exactly twice and the subject chooses Route 2, their regret will be 0 – (-1250) = 1250, because the optimum choice was Route 4 with a payout of zero. Another key note to make about this payout schedule is that the payout does not redistribute after each selection. The columns can be viewed as a stack where each payout choice remains at the top until chosen. For example, from the schedule below in Figure 3, if a subject does not choose Route 1 until trial 6, their payout would still be 100.

Route 1	Route 2	Route 3	Route 4	Subject's Selection	Regret
100	100	50	50	Trial 1: Route 4	100 - 50 = 50
-350	0	-50	-250	Trial 2: Route 4	100 - (-250) = 350
-250	-1250	-50	0	Trial 3: Route 2	100 - 100 = 0
0	0	0	0		
-200	0	-50	0		
0	0	0	0		
-300	0	-50	0		

Figure 3. Payout schedule excerpt. The blue cell indicates the optimal decision; the yellow cell shows the subject's selection on trial 1; the green cell indicates the subject's selection on trial 2.

D. THESIS GOALS

This thesis has four objectives: (1) find a threshold that delineates between high and low regret (decision performance), (2) combine the decision performance data with the cognitive state data, (3) validate these results and CAPTTIM, and (4) develop a visualization method for displaying a subject's CAPTTIM category on a trial-by-trial basis. A superficial analysis of regret, from the previously collected data, showed that it was consistent with subject's actual performance, as measured by total damage score. Subjects that identified the convoy route with the optimal long term result had a decreasing amount of regret

(Nesbitt et al., 2015). If a threshold for regret is validated, then the utility of CAPTTIM can be tested with other military tasks. CAPTTIM has the potential to provide the instructor with real time guidance on type and timing of intervention in a training scenario.

THIS PAGE INTENTIONALLY LEFT BLANK

II. METHODS

The data used in the analysis portion of this thesis was previously collected from the convoy task and de-identified. This chapter will list in detail the tools and methods used to analyze the regret data in an effort to delineate a threshold between high and low regret. These methods were initially tested (i.e., piloted) on a randomly selected subset of eight of the 34 participants who completed the convoy task. Data from the remaining 26 participants would be used to test the final, selected method. An iterative process was conducted to find an appropriate method, in which initially selected methods informed and directed the subsequent methods. As a result, all the methods described below are more or less in chronological order (exponentially weighted moving average, simple moving average, x bar control chart, change point analysis).

A. STATISTICAL SOFTWARE: R STUDIO

The programming language R (R Development Core Team, 2008), which was developed for statistical computing, was utilized for the analysis of the regret data collected from the convoy task (Nesbitt et al., 2013). All the code written for this analysis can be viewed in Appendix B. R-Studio, the integrated development environment (IDE) that was developed for the R language, was used to develop the code that analyzed the regret data. R-Studio is an open source IDE that allows the user to code line by line the exact code for statistics equations. R-Studio varies from a statistics program like JMP in that it requires the user to understand and program every function rather than operating in a drag and drop type fashion like JMP.

B. METHODS USED TO DELINEATE HIGH AND LOW REGRET

Each of the following methods used to research a threshold delineating between high and low regret were coded and calculated in R Studio. Once an analysis was conducted with a specific method, the research team was briefed on the results. This collaboration led to the rejection of three of the four methods utilized to distinguish a regret threshold.

The following sections will chronologically list each of the four methods that were researched. A thorough explanation of each method and how it was used in an attempt to delineate between high and low regret will be given. Additionally, the shortfalls of the first three methods to delineate between high and low regret will be explained.

1. Exponentially Weighted Moving Average (EWMA)

The following section will give a brief introduction of the EWMA equation and its common uses. The next section will discuss how the EWMA was used to analyze the data collected for this thesis. This was the first method explored in an effort to find a threshold to delineate decision performance (high versus low regret).

a. Explanation of EWMA Equation and Uses

"The Exponential Weighted Moving Average (EWMA) chart is used for monitoring process by averaging the data in a way that give less weight to old data as samples are taken and gives more weight to most recent data" (Braimah, Osanaiye, Omaku, Saheed, and Eshimokhai, 2014, p. 1). EWMA also is very effective at detecting minor changes in the process mean (Braimah et al., 2014). It was originally developed by S. W. Roberts in 1959 as a means of monitoring control/performance charts in industrial processes (Braimah et al., 2014). It also has been very useful in time series analysis and forecasting (Braimah et al., 2014). The following is how an individual EWMA value is calculated as

$$Z_i = \lambda X_i + (1 - \lambda) Z_{i-1} \text{ ,}$$

where Z_i is the EWMA control statistic, λ is the weighted parameter, and X_i is the actual observed data value

A key difference between EWMA and a simple moving average is that EWMA considers all previous data points, while a simple moving average only considers data points within a specified window (Braimah et al., 2014). “EWMA weights samples in geometrically decreasing order so that the most recent samples are weighted most highly while the most distant samples contribute very little” (Braimah et al., 2014, p. 2). This weighted parameter, λ ($0 < \lambda \leq 1$), is a mathematical representation of how heavily memory of past data is relied upon (Kalgonda, Koshti, and Ashokan, 2011). As λ increases from zero to one, more weight is placed on recent data points and less weight is placed on distant data points. If $\lambda = 1$, then 100 percent of the weight is placed on the most recent data point and no weight is placed on the past (Kalgonda et al., 2011). The sensitivity of the EWMA to small shifts in the process mean is reliant upon the value of λ (Kalgonda et al., 2011).

The use of EWMA as a means of detecting changes in regret was based on the EWMA’s sensitivity to small shifts and reliance on memory. Because decisions on the convoy task rely heavily upon working memory and the influence of past decisions on future decisions (Kennedy et al., 2013), this method of averaging regret seemed more appropriate than a simple moving average.

Using EWMA to analyze regret was the initial approach taken because it worked exceptionally well in characterizing subject’s cognitive state based on decision time latencies in the convoy task. An effective threshold delineating between the cognitive states of exploration and exploitation was applied to this EWMA and accurately portrayed subject’s transition between these two states.

The threshold that was used was double the standard deviation of each subject's latency times in decisions that resulted in low damage. The EWMA equation for time latency utilized a λ value of 0.1. This λ value means that subjects had a heavy reliance on past decisions, since $(1 - \lambda)$ determines the weight placed on past data points. This code was modified to analyze regret and utilized the same value of λ.

b. EWMA of Regret

The initial EWMA of regret looked at the mean values of regret. This meant that the EWMA was looking at the cumulative regret divided by the number of trials. This analysis produced some interesting results. However, upon further discussion with the research team and additional analysis, the use of the mean regret as the values on which to conduct the EWMA was determined to be incorrect. By using mean regret the values were essentially being smoothed twice. Dividing the cumulative regret by the trial was taking an average after every trial; this average was again being averaged with the EWMA based on the weight placed on past data. This realization led to the decision that the EWMA should be conducted on the regret per trial for each subject.

By using the regret received by the subject at each trial, the EWMA was looking at actual values and not an already averaged value. The result was much more volatile changes in the EWMA.

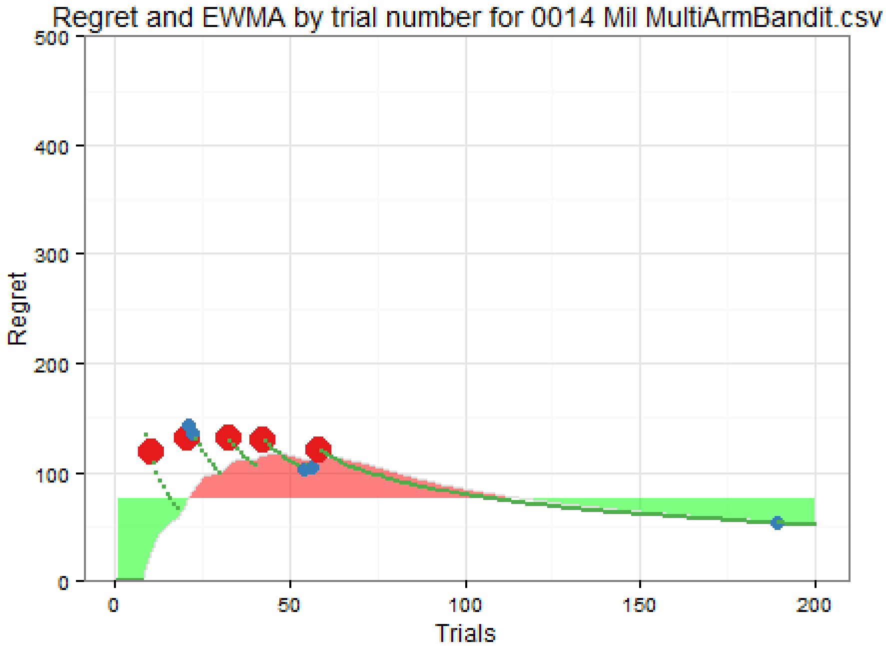


Figure 4. EWMA of regret for Subject 14 using mean regret. Mean regret proved to be inappropriate as it was performing a EWMA on an already averaged regret value. This accounted for the much less volatile spikes in regret value. The large red dots are high damage instances. The medium blue dots are medium damages, and the small green dots are low damage instances. The shaded red area is where the EWMA is above the threshold and the shaded green area is where the EWMA is below the threshold. The threshold is calculated as 0.5 times the standard deviation of the mean regret.

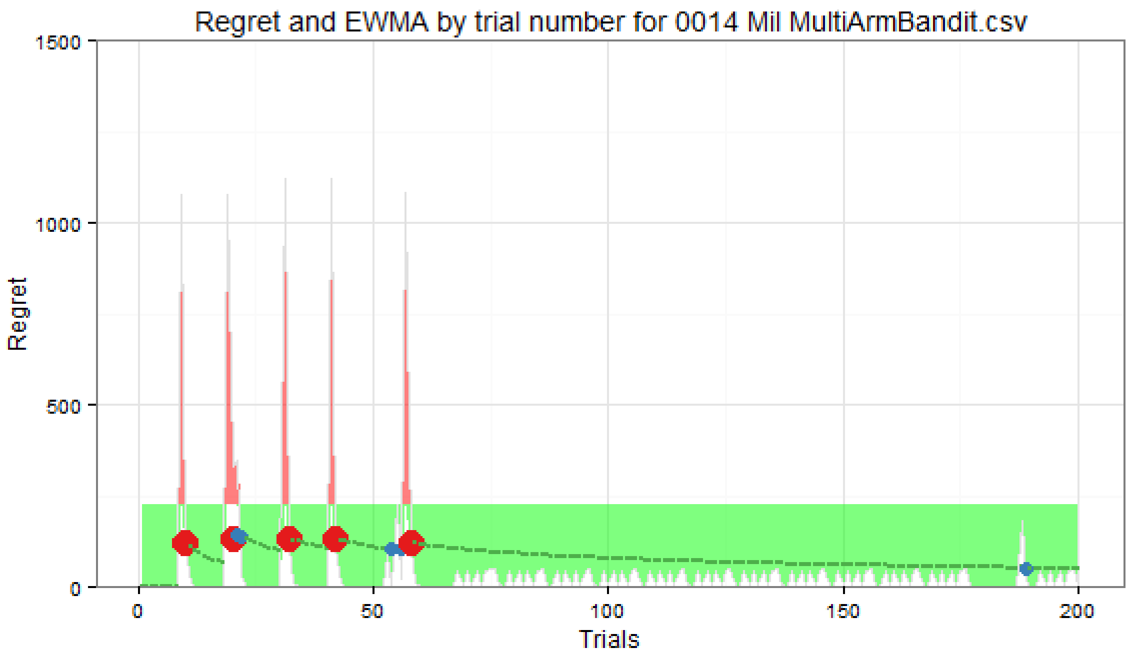


Figure 5. EWMA of regret for Subject 14 using regret received at each trial. The volatility in high regret is seen with the sharp red peaks which is where regret reaches values of 1250 for high friendly damage. The red, blue, and green dots are for high, medium, and low damages respectively. These dots are plotted along the mean regret line. Shaded red areas are above the threshold, while shaded green areas are below the threshold. The threshold is defined as the standard deviation of the regret received per trial.

The threshold value for the EWMA conducted on mean regret had to be adjusted to one half the standard deviation of regret in order to have the EWMA fall above and below the threshold, as can be seen in Figure 4. This adjustment was as a result of averaging an already averaged value. The threshold for the EWMA conducted on regret received per trial was strictly the standard deviation of the regret per trial and did not require any fractional adjustment. After discussion and further analysis with the research team, it was suggested that a sensitivity analysis of λ to the regret per trial data be conducted. Based on the sensitivity analysis the ability to tune λ to the actual data could be achieved.

This sensitivity analysis of regret per trial to λ resulted in the realization of the difficulty of tuning this parameter for this use case. The analysis showed that a λ value of 0.9 achieved the line of best fit for each subject to the actual regret data (this realization is trivial given the EWMA equation). This value of λ illustrated that subjects placed very little weight on past regret and that the immediate results influenced their decision the most. Figure 5 illustrates this point —had Subject 14 weighted past decisions heavily, the spikes in regret would have become less volatile and been spread across future decisions, illustrating that he/she had been influenced by the previous decision.

Thus, this EWMA was fit to the actual regret per trial data and led to highly volatile changes in regret. Despite a defined payout schedule, values of regret are very random across subjects with a wide range of possible values. For example, one subject may have only experienced regret values of 50 if they converged on the optimal path, while another subject may have experienced regret values of 1250 since they did not converge on the optimal path. The high volatility of these values made defining a single threshold difficult, since regret could range from 0 to 1250. This issue made it difficult to classify into which category of the CAPTTIM model a subject should be categorized. Therefore, other approaches were sought. The next method examined was the simple moving average.

2. Simple Moving Average

Rather than looking at a trial by trial analysis of whether regret was increasing or decreasing, a simple moving average was conducted to "block" regret by a specific number of trials. As a reminder, simple moving average differs from EWMA in that it only considers the data within a specific window, whereas the EWMA considers all data points and weights them according to the value of λ. Two approaches were taken: (1) the simple moving average looked at a moving window of five trials throughout the 200 trials of regret data (2) the simple moving average did the exact same calculation with a moving window of

10 trials. The moving window of five trials allowed for more granularity in observing this subject's changes in regret. Utilizing a larger window gives less blocks to analyze changes in regret and thus does not provide as much sensitivity for changes in regret (see Figures 6 and 7). As a result, the simple moving average that utilized a window of 5 trials was used for the follow on analysis of regret.

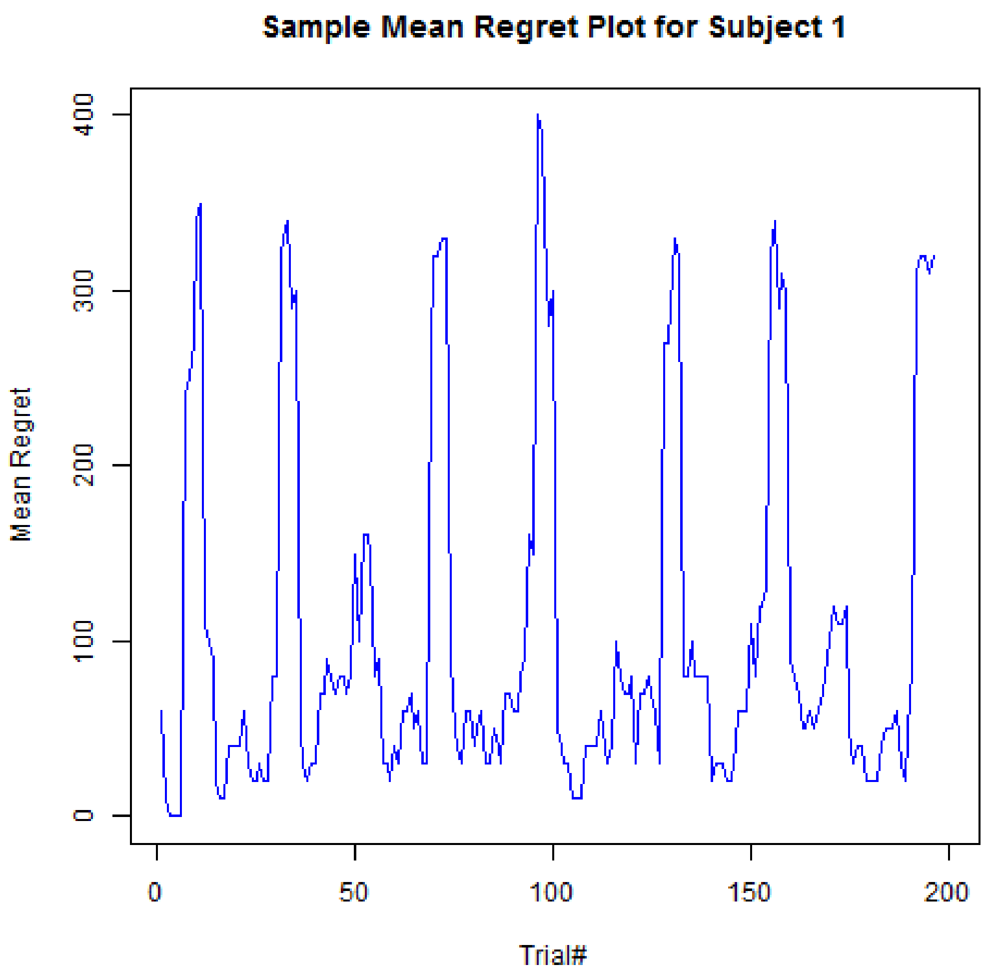


Figure 6. Simple moving average of regret per trial for Subject 1 with a window of 5 trials. The solid blue line shows the averaged regret and how high values in regret influenced the average for the 4 previous and 4 successive trials. Had a simple moving average not been used, high values of regret would be single vertical lines.

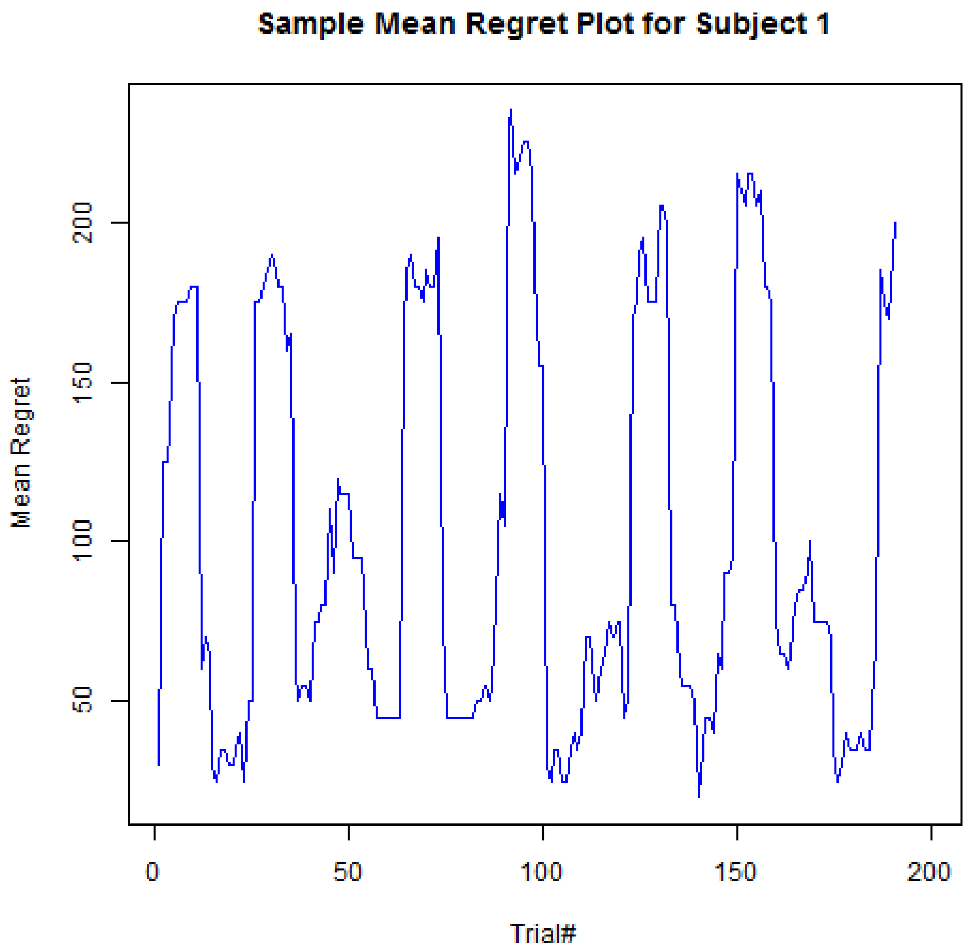


Figure 7. Simple moving average of regret per trial for Subject 1 with a window of 10 trials. The solid blue line shows the averaged regret and how high values in regret influenced the average for the 9 previous and 9 successive trials. Comparison to Figure 6 shows how, for the same subject, the spikes in high regret are broadened by utilizing a larger window.

The use of a simple moving average of regret provided more insight into defining a subject's performance than the EWMA of regret. Because regret for most subjects was extremely random, trying to define a threshold to differentiate between high and low regret using an EWMA was very difficult to do. The simple moving average allowed an analysis of discrete blocks to determine the slope of the line, which in turn showed whether regret was increasing or decreasing at

specific points. However, as described in the section below, it was discovered that the simple moving average method also had drawbacks.

3. X-Bar Control Chart

Instead of looking at a simple moving average of regret and applying a threshold that delineated between high and low regret, a better approach could be to create a control chart that defines a median and an upper control limit. As long as the value falls within the upper control limit, the subject is deemed within tolerance or having low regret. The control chart made it a lot easier to classify subjects into their specific category in CAPTTIM. Originally the control chart looked at using the mean of regret per trial plus the standard deviation of regret to define the upper control limit. This upper control limit adjusted utilizing the same 5 trial window that the simple moving average utilized. However, what the research team found was that the mean was not a useful metric for determining the upper control limit of the control chart. This was due to the fact that regret has possible values ranging from 0 to 1250. With such volatility in values, the mean and standard deviation are skewed due to these high spikes in regret experienced by most subjects. Therefore, the upper control limit was falsely classifying subject performance, and as a result very few subjects were being classified as out of tolerance (high regret). In fact, most subjects were being classified as having low regret despite their actual overall performance (final damage score). A histogram of regret was created, in order to illustrate the unsymmetrical characteristic of the regret data (see Figure 8).

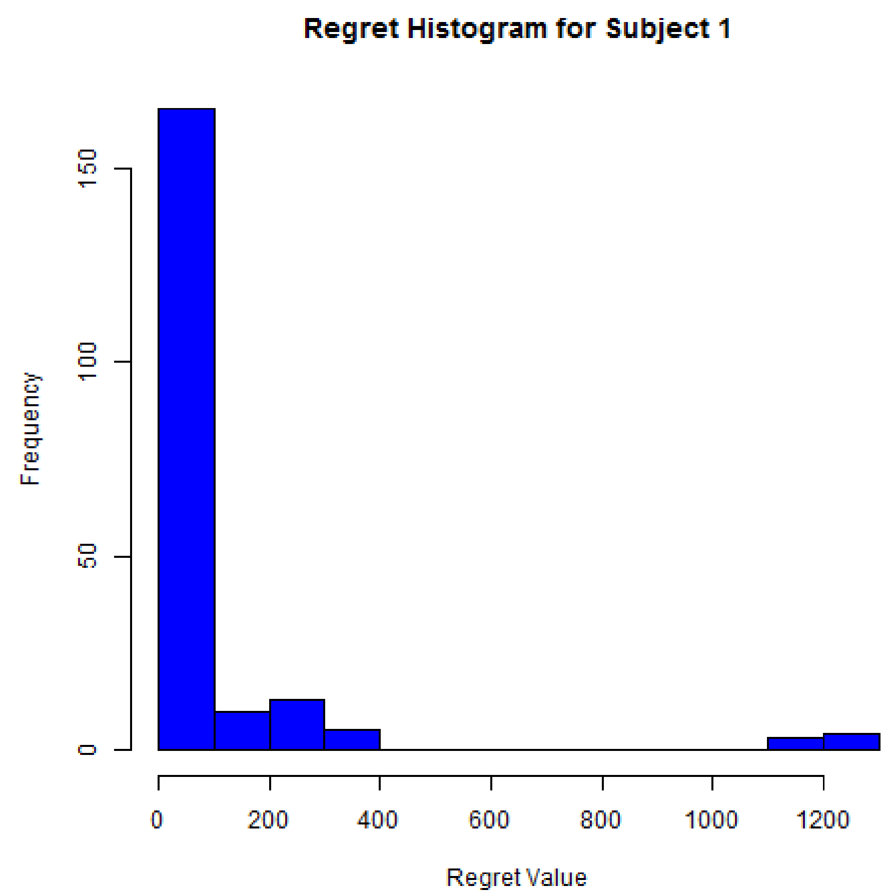


Figure 8. Histogram of regret data for Subject 1. This clearly illustrates that the majority of regret values experienced by Subject 1 are of magnitude 50 and that the high spikes in regret only occurred a handful of times.

Due to the variation in the data for regret, the next approach taken was to look at the median of regret versus the mean. Additionally the research team recommended looking at a window of 20 trials to calculate the median and upper control limit in order to provide a more stable analysis of tolerance. This window of 20 trials was chosen based on the payout schedule and when these large values of regret were incurred. Additionally the window of 20 trials provided an appropriate window in which subjects would be allowed to illustrate reinforcement learning and make mistakes and adjust their course of action. Smaller windows proved to be too restrictive and classify subjects out of tolerance too hastily. The

new upper control limit for the X-Bar chart was then calculated as the median plus the median absolute deviation for the moving window of 20 trials. Figure 9 shows the X Bar control chart for Subject 1. The solid blue line is the simple moving average described before, and the dashed red line is the median plus the median absolute deviation, which is recalculated every 20 trials. Points on the simple moving average that were above the dashed red line are considered out of tolerance (high regret), while points below the red dashed line were considered within tolerance (low regret) (see Figure 9).

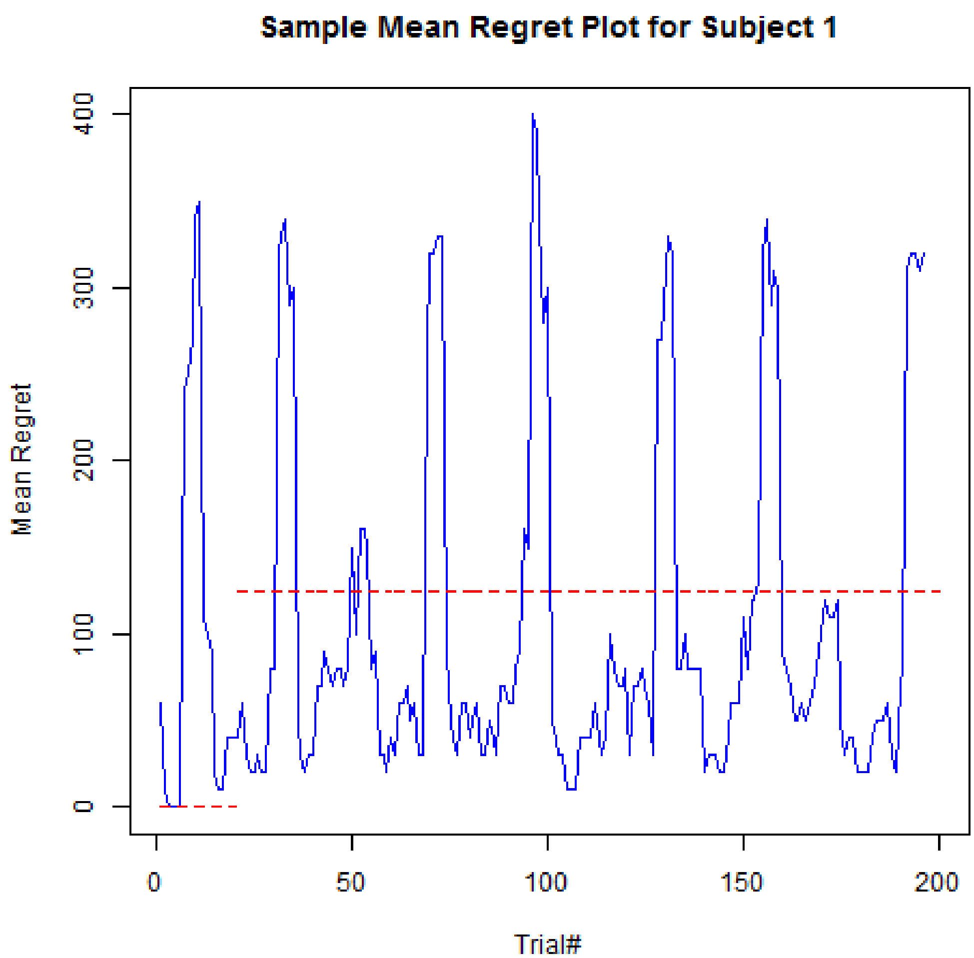


Figure 9. X-Bar control chart for Subject 1. The solid blue line is the simple moving average that was previously discussed. The dashed red line is the upper control limit. The upper control limit is defined as the median plus the median absolute deviation and

is recalculated every 20 trials.

4. Change Point Analysis

After discussion with the research team and a recommendation from the team statistician, Dr. Fricker, a change point analysis was conducted to determine the best window size of trials to create the upper control limit for the X-Bar control chart. Change point analysis is useful in determining if a change occurred, how many changes occurred, when the changes occurred, and provides with what confidence the changes occurred (Taylor, 2000). Change point analysis is extremely flexible and can be performed on all types of time ordered data to include, attribute data, non-normal distributions, ill-behaved data, and data with outliers (Wayne, 2000). A key difference between change point analysis and control charts in the context of regret is that control charts can be generated following each individual trial, while a change point analysis can only be generated retrospectively (Wayne, 2000). Change point analysis is typically more sensitive and can often detect changes in the process mean that are missed by the control chart, thus the two methods are best employed in a complimentary fashion (Wayne, 2000).

5. Final Method: Combination of Control Chart and Change Point Analysis

Combining control chart and change point analysis, in this complimentary fashion, is the method being employed in this thesis. The statistical computation language R contains built in packages for conducting change point analysis. The R package utilized in this analysis was the segment neighborhood (SegNeigh) algorithm (Killick, & Eckley, 2014). This algorithm utilizes dynamic programming to calculate the optimal segmentation for m + 1 change points and reuses the data calculated for m change points (Killick et al., 2014). This essentially means, that the algorithm searches over all previous change points and chooses the one that results in the optimal segmentation up to that time (Maidstone, Fearnhead, & Letchford, 2013). This package takes a variable Q that specifies the maximum

number of change points to identify. This was useful in the analysis of the non-normal data contained in the data set of regret per trial. Due to the volatility of the regret per trial data, running a change point analysis package that identified every change point was not useful. However, by specifying a smaller number of change points (Q=15) the analysis was able to yield results that were useful in delineating between high and low regret. Figure 10 shows the change point analysis performed on Subject 1.

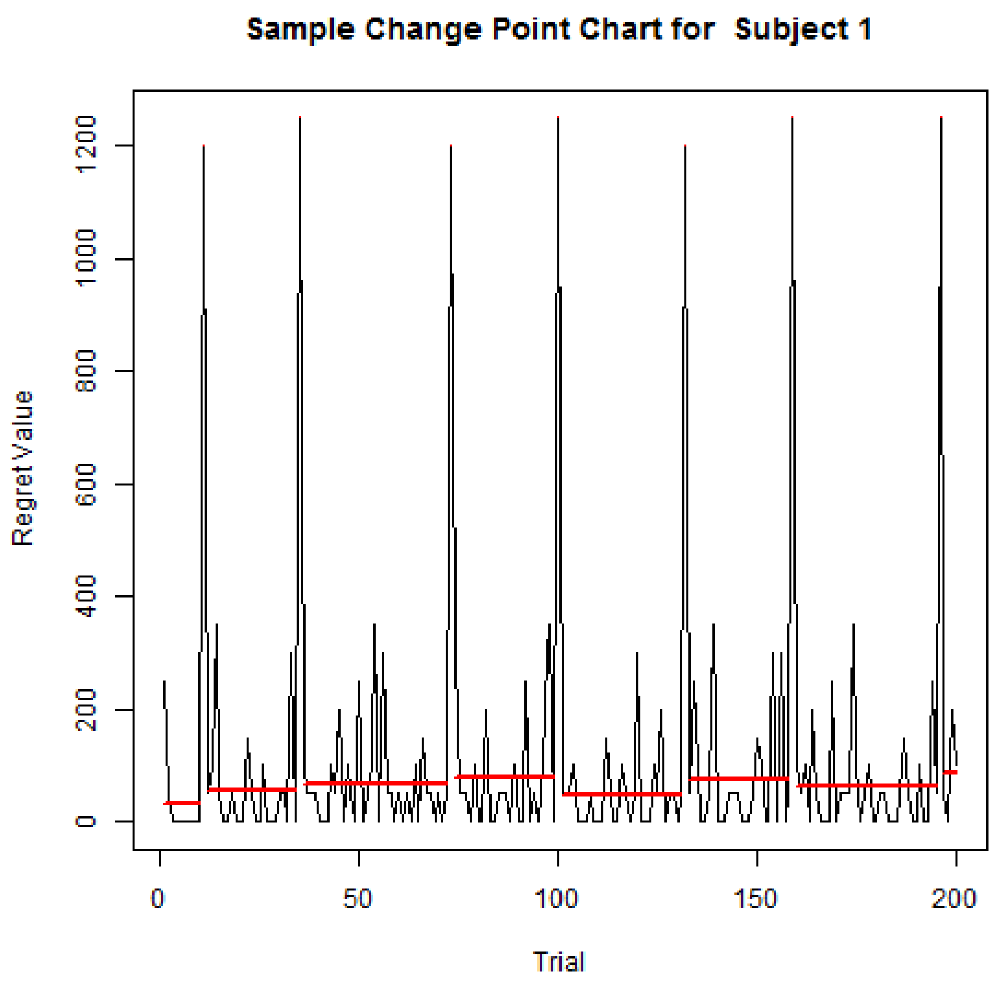


Figure 10. Change point analysis for Subject 1. The solid black line is the regret per trial data. The solid red lines are the process means returned by the change point analysis—they represent the process mean for that range of trails. The large spikes in regret incurred a change in the process mean that spanned the single trial in which the regret was incurred.

After studying the change point analysis and further discussion with the research team, it was decided that, rather than using an X-Bar control chart, creating a box plot of the means associated with each change point and determining if the mean was above or below the median would accurately delineate between high and low regret. Because the change point analysis returns the mean as well as the trial number for each change point, the subject can be accurately categorized in CAPTTIM for a range of trials. This was the final method decided upon for analyzing regret for the subset of 8 subjects along with the subsequent 26 subjects.

In addition to the use of the change point analysis to delineate between high and low regret, the research team decided to add an additional metric for determining decision performance. Subjects that chose route 1 or 2 after trial 100 would be automatically classified as having high regret. This metric took into account the time and duration of the experiment and at which point the optimal performers converged on the ideal decision.

THIS PAGE INTENTIONALLY LEFT BLANK

III. RESULTS

By conducting the change point analysis on all 34 subjects and comparing the resulting means with the median off all change point means, an effective threshold for delineating between high and low regret was established. Once the threshold for delineating between high and low regret was obtained, the data could then be compared with the cognitive state of the subject in order to categorize them in CAPTTIM. This section will detail how each subject's regret was categorized and then compared with the cognitive state data.

A. OVERVIEW OF COGNITIVE STATE DATA DEVELOPED FROM PRIOR RESEARCH

A subject's cognitive state was previously categorized by Maj Pete Nesbitt, who utilized an EWMA of the latency in decision-making times. A threshold was then applied to the EWMA in order to delineate between the cognitive states of exploration and exploitation. The threshold that was utilized was two times the standard deviation of latency in decision-making times immediately following trials that resulted in low damage. It was assumed that decision times after receiving low damage would be relatively fast, and therefore, could be used to determine an individual subject's baseline latency time. In contrast, it was assumed that decision times following trials that resulted in high or medium damage would be longer, because subjects typically reflected on the negative feedback. The threshold was specific to each subject since it was calculated using their baseline. This threshold accurately delineated between exploration and exploitation for all 34 subjects. This prior work allowed the research team to know on a trial-by-trial basis whether the subject was exploring or exploiting (see Figure 11). This knowledge was crucial in the development of the CAPTTIM categorization algorithm.

Most subjects illustrated a pattern of taking longer to make decisions in the beginning of the convoy task when they were exploring and gathering information on the environment (higher latency times between decisions). Most subjects then transitioned to making decisions more rapidly (lower latency times between decisions) once they believed that they had converged on the correct choice and were exploiting that path. This pattern can easily be seen in Figure 11, where Subject 4 spent approximately 45 trials exploring (shaded orange region) and then transitioned to exploitation (shaded blue region) from trial 45 to 200. As can be seen from Figure 11, even though Subject 4 began exploiting the decision that he/she thought was the correct decision, heavy friendly damages (large red dots) were incurred throughout the remainder of the trials. Because Subject 4 incurred heavy and medium friendly damages throughout the 200 trials, his/her final damage score was much lower than those of subjects who converged on the optimal choice. As a reminder, each subject began the experiment with a positive final damage score of 2000. When they received friendly damage this would deduct from their final damage score and when they inflicted damage on the enemy this would increase their score. The average final damage score across all 34 subjects was 2,402.94. Subject 4's final damage score was 2050.

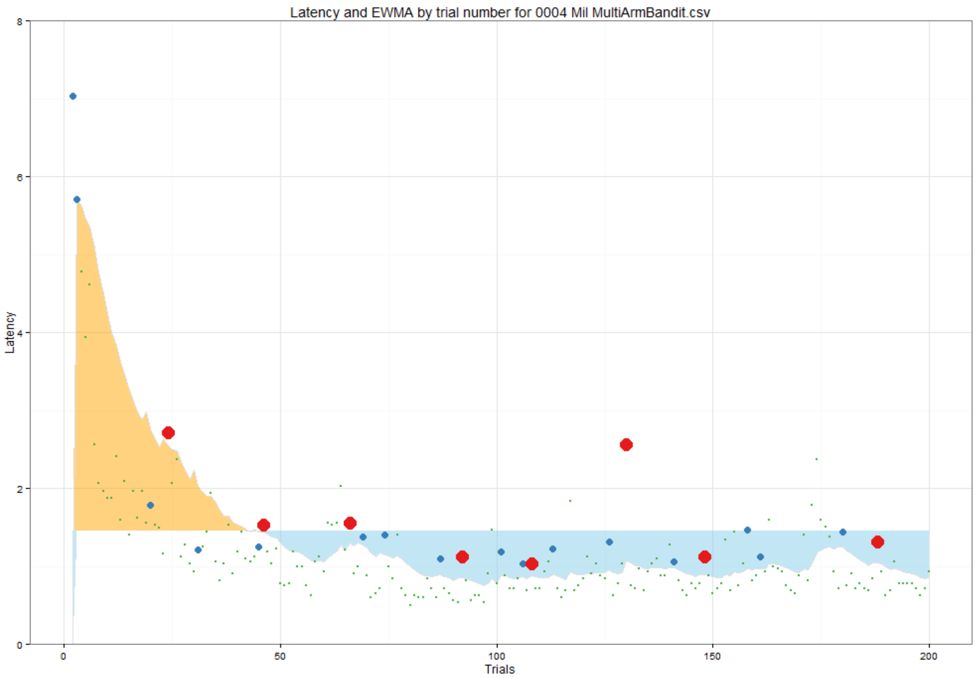


Figure 11. EWMA of latency in decision-making times for Subject 4. The y-axis is latency in decision-making times and the x-axis is the number of trials. The colored dots represent damage incurred and are plotted at the actual latency in decision-making time versus the EWMA. The color and size of the dot is correlated with the level of damage incurred on the preceding trial. Red dots are high damage, blue dots are medium damage, and green dots are low damage. The orange shaded regions are where the EWMA is above the threshold (exploration) and the blue shaded regions are where the EWMA is below the threshold (exploitation).

The following example is of a subject who illustrated optimal exploration of the environment followed by exploitation of the optimal choice. Figure 12 is the EWMA of latency in decision-making times for Subject 14. Subject 14 followed the typical pattern observed for most subjects, by exploring in the beginning (shaded orange region) and then transitioned to exploiting (shaded blue region). Subject 14 transitioned between exploration and exploitation by approximately trial 30. While Subject 14 took some medium damages (medium blue dots) and high damages (large red dots) in the beginning of his/her exploitation phase, he/she eventually converged on the optimal decision and incurred very little

damage throughout the remaining trials. As a result, Subject 14's final damage score was 4700 compared to Subject 4's score of 2050.

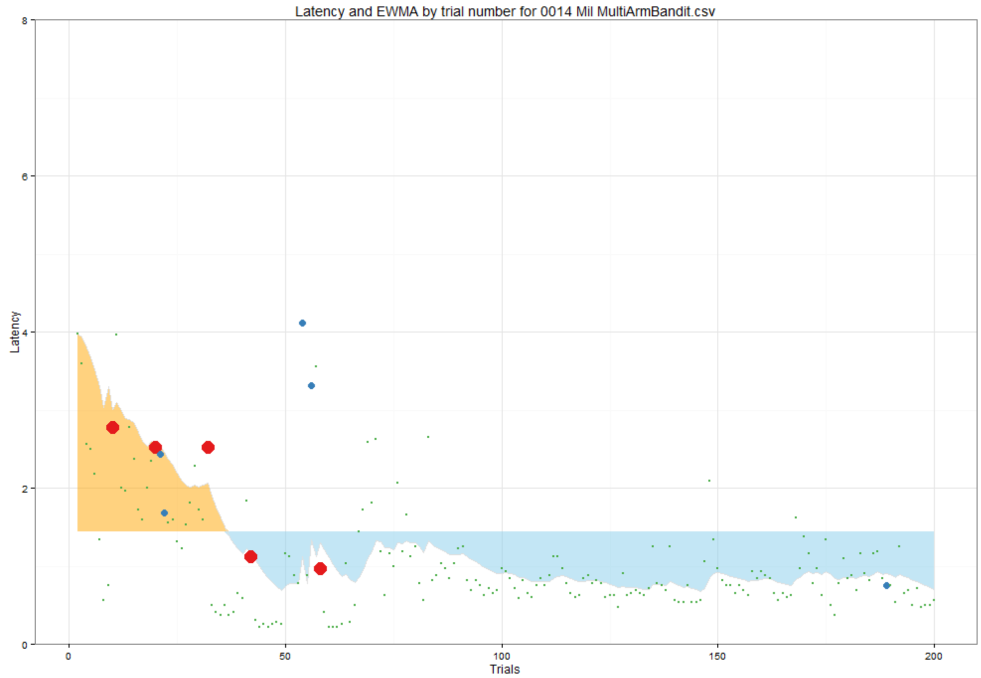


Figure 12. EWMA of latency in decision-making times for Subject 14. The y-axis is latency in decision-making times and the x-axis is the number of trials. The colored dots represent damage incurred and are plotted at the actual latency in decision-making time versus the EWMA. The color and size of the dot is correlated with the level of damage incurred on the previous trial. Red dots are high damage, blue dots are medium damage, and green dots are low damage. The orange shaded regions are where the EWMA is above the threshold (exploration) and the blue shaded regions are where the EWMA is below the threshold (exploitation).

These examples demonstrate that knowing a subject's cognitive state does not provide sufficient insight into their actual decision performance. Subjects 4 and 14 showed similar cognitive state patterns yet had very different decision performance. Thus, the next step was to combine the subject's cognitive states with the categorization of their actual performance (high versus low regret), which was the focus of the research conducted in this thesis.

B. CHANGE POINT ANALYSIS COMBINED WITH COGNITIVE STATE DATA

The cognitive state data from above was then taken and combined with the change point analysis data that delineated between high and low regret. This delineation provided a metric to gauge a subject's actual performance. The combination of actual decision-making performance with cognitive state allowed for the categorization of subjects into CAPTTIM.

1. Delineating High and Low Regret Using Change Point Analysis

Using the change point analysis data, subjects were categorized as having high or low regret on a trial-by-trial basis. The change point analysis returned 15 change points for each of the 34 subjects. These change points represent instances where a subject's process mean changed. The reason that 15 change points were returned was as a result of the method used within R (SegNeigh) to conduct the change point analysis. The number of change points was limited to 15, due to the volatility of the regret data. Regret per trial values vary between 0 and 1250 with intermediate values of 100, 200 and 300. By limiting the number of change points the significant changes were readily identified, while the minor changes were allowed to occur without changing the process mean. If every change point were identified the number of change points would have been too numerous to provide any use for analysis. The change point and its associated process mean were then compared with the median of all 15 process means. This comparison looked at windows of trials on the basis of the process means returned from the change point analysis (see Figure 13). The process mean for that window of trials was then compared with the median of the process means to determine whether it fell above or below the median. If the process mean was above the median, the subject was categorized as having high regret; if the process mean was below the median, the subject was categorized as having low regret. Figure 13 clearly indicates that Subject 4 experienced peaks of high regret throughout his/her 200 trials, which resulted in a much lower final damage score.

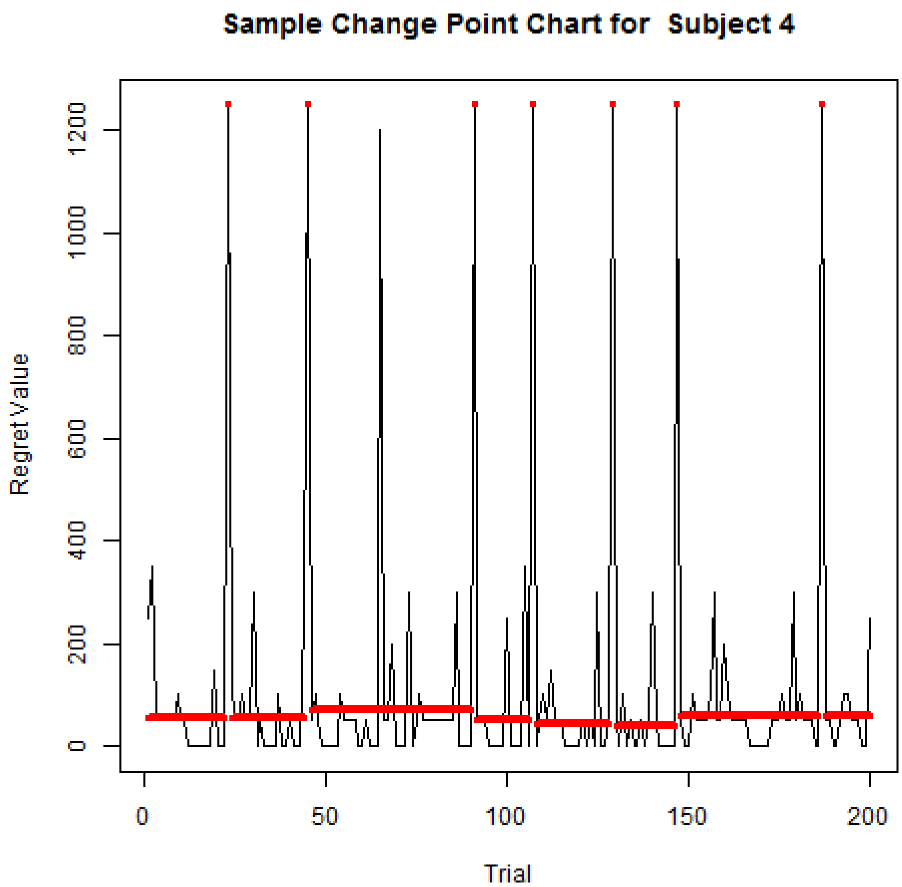


Figure 13. Change point analysis for Subject 4. The y-axis is the regret per trial value, while the x-axis is the trial number. The red lines are the process means returned from the change point analysis. The spikes in the regret value are a result of the subject receiving heavy friendly damage and incurring high regret. These spikes result in a change point that exists over just one trial. The other, longer red lines are where the process mean did not change for that range of trials.

The following information illustrates the change point analysis results for a subject who converged on the optimal choice. Figure 14 is the change point analysis chart for Subject 14. Subject 14 clearly illustrated the ideal exploration phase where heavy damage is expected and encouraged in order for the subject to fully explore the environment and identify the optimal choice. This exploration phase was followed by an ideal exploitation phase, where Subject 14

experienced minimal regret. Because Subject 14 experienced minor regret for the majority of trials, his/her final damage score was much higher than that of Subject 4 (4700 vs. 2050). Another interesting point illustrated by Subject 14, was that he/she experienced numerous change points in the beginning, but after trial 60 (approximately) the process mean remained constant.

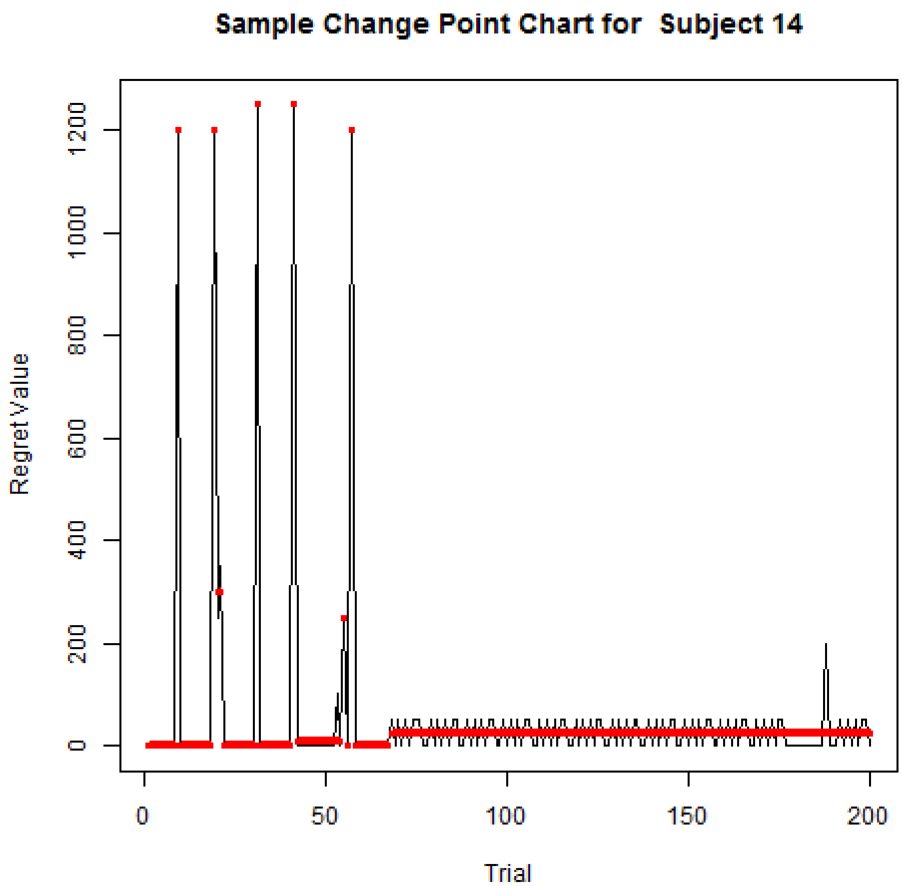


Figure 14. Change point analysis for Subject 14. The y-axis is the regret per trial value, while the x-axis is the trial number. The red lines are the process means returned from the change point analysis. The spikes in the regret value are a result of the subject receiving heavy friendly damage and incurring high regret. These spikes result in a change point that exists over just one trial. The other, longer red lines are where the process mean did not change for that range of trials.

Once a threshold was established that effectively delineated between high and low regret and provided a method for gauging actual decision performance, the research team had all the requisite information required for categorizing subjects within CAPTTIM. This ability to categorize subjects within CAPTTIM fulfilled a primary goal of this thesis.

2. Combining Cognitive State and Decision Performance to Categorize Subjects within CAPTTIM

The combined cognitive state data and decision performance data allowed for the categorization of subjects within CAPTTIM to be accomplished. Figure 15 shows the CAPTTIM categorization algorithm used to properly assign subjects within their appropriate category.

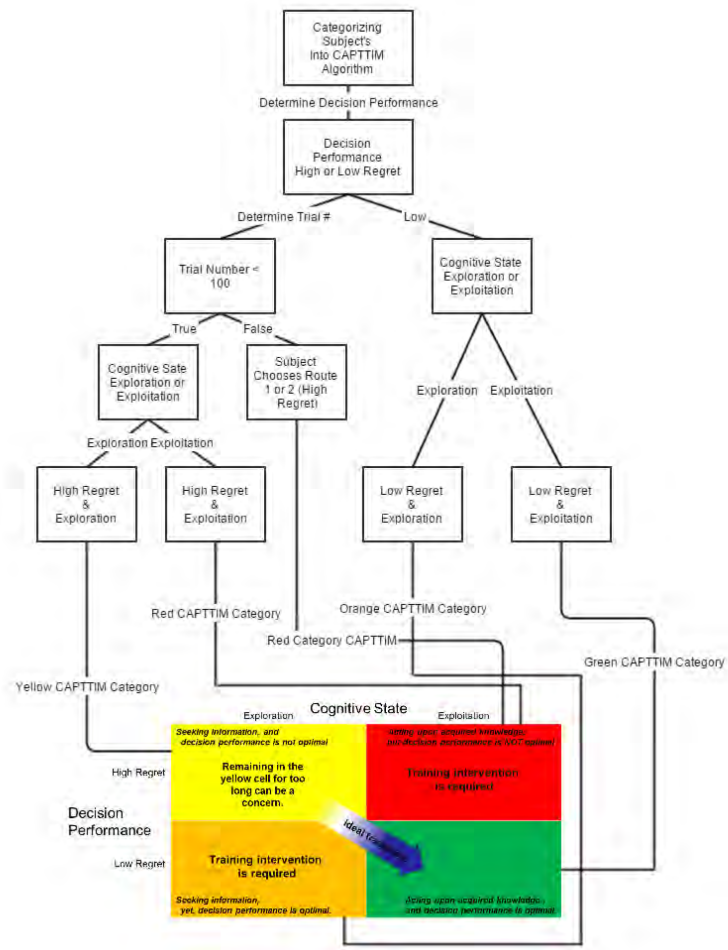


Figure 15. CAPTTIM categorization algorithm. This figure illustrates how each subject is categorized in CAPTTIM based on decision-making performance (measured by regret) and cognitive state (measured by latency in decision-making times).

Because the change point analysis of regret and EWMA of latency in decision-making times delineate between decision performance and cognitive state for a range of trials, a graphical representation was developed that represents what category of CAPTTIM a subject is in on a trial by trial basis. This representation was overlaid on the regret per trial graph in order to illustrate how CAPTTIM could be used to provide instructors information on type and timing of intervention.

Figure 16 is the CAPTTIM categorization chart for Subject 4. Figure 16 clearly shows that Subject 4 experienced high regret at times during his/her exploration phase (yellow block), but never fully explored the entire environment (orange blocks). After a brief exploration phase (approximately 45 trials), Subject 4 transitioned to the exploitation phase. For windows of trials Subject 4 exploited decisions that resulted in low regret (green blocks). However, these windows were often interrupted by exploited decisions that resulted in high regret (red blocks). These repeated exploited decisions with high regret were a clear indicator that Subject 4 did not converge on the optimal choice.

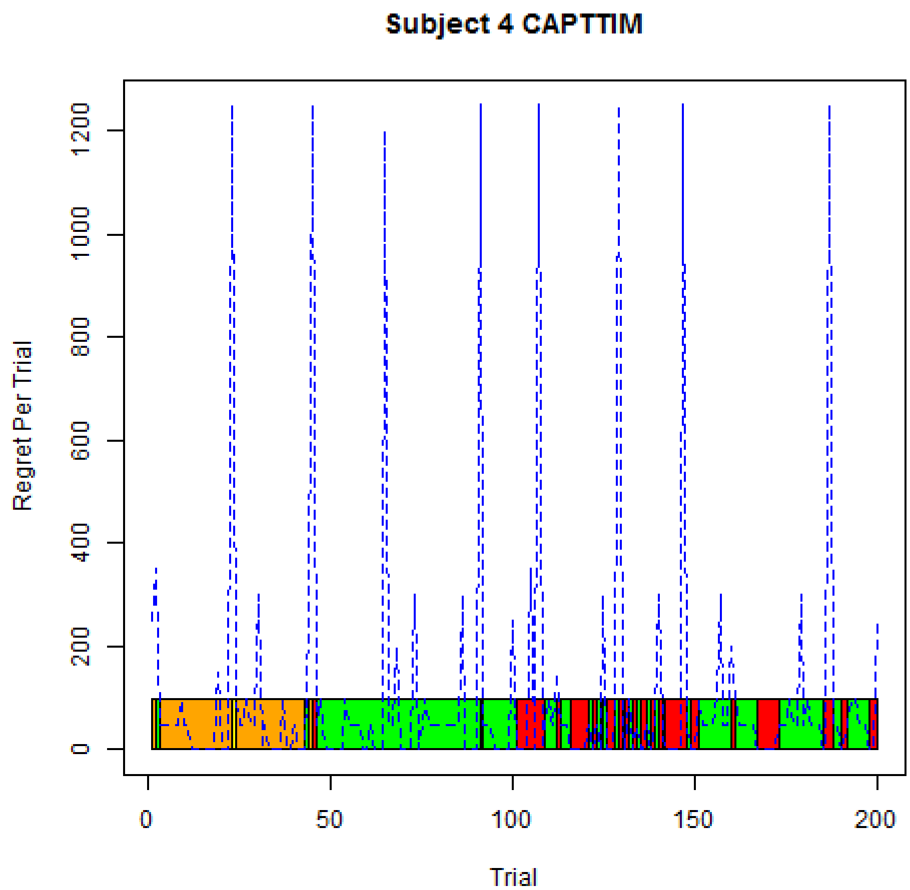


Figure 16. CAPTTIM categorization chart for Subject 4. The color-coded bar at the bottom of the chart correlates to the category color found within the CAPTTIM model. Yellow is high regret and exploration. Orange is low regret and exploration. Red is high regret and exploitation. Green is low regret and exploitation.

Figure 17 is the CAPTTIM categorization chart for Subject 14. This figure accurately portrays that Subject 14 experienced high and low regret during his/her exploration phase (yellow and orange blocks), and even experienced a couple of poor choices during the initial exploitation phase (red blocks). For the vast majority of trials, however, Subject 14 made the ideal transition and converged on the optimal choice (green block) and did not deviate from the optimal choice for the remaining trials.

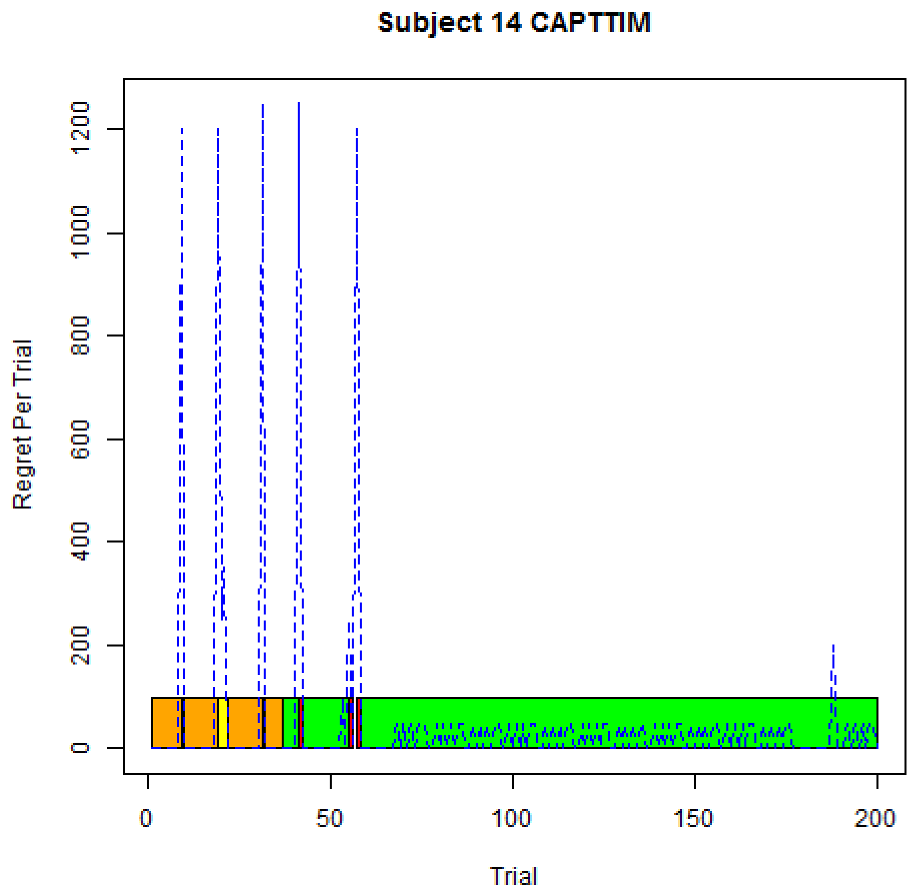


Figure 17. CAPTTIM categorization chart for Subject 14. The color-coded bar at the bottom of the chart correlates to the category color found within the CAPTTIM model. Yellow is high regret and exploration. Orange is low regret and exploration. Red is high regret and exploitation. Green is low regret and exploitation.

The CAPTTIM categorization charts for Subjects 4 and 14 clearly illustrated typical patterns observed across the 34 subjects. Subject 4 illustrated

how the optimal path was never identified and exploited. This decision pattern would have resulted in an instructor intervention based on the CAPTTIM results. Subject 14, however, converged on the optimal choice and exploited. Thus, this decision pattern would have resulted in no instructor intervention being needed. The research team observed that the subjects fell into three typical groups consisting of (1) subjects who explored and eventually identified the optimal choice (n = 9), (2) those who explored and exploited non-optimal choices (n = 21), and (3) subjects who never transitioned from the exploration cognitive state to the exploitation cognitive state (n = 4). This third group would have required instructor intervention, which was accurately identified using the CAPTTIM categorization charts. This third group is illustrated by subject 11 in Figures 18 and 19.

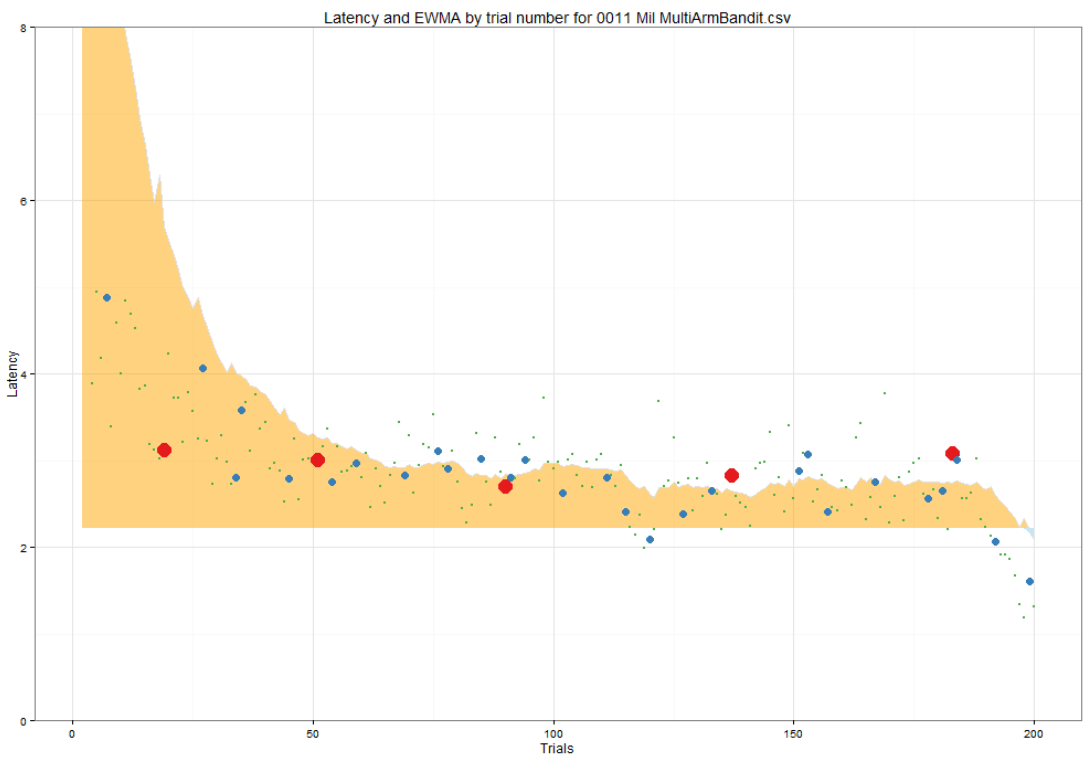


Figure 18. EWMA of latency in decision-making times for Subject 11. The x- and y-axis are the same as the previously described graphs. Note that Subject 11's EWMA of latency in decision-making times never falls below his/her threshold (shaded orange region). This subject spent the entire time exploring the environment and never exploited any decisions.

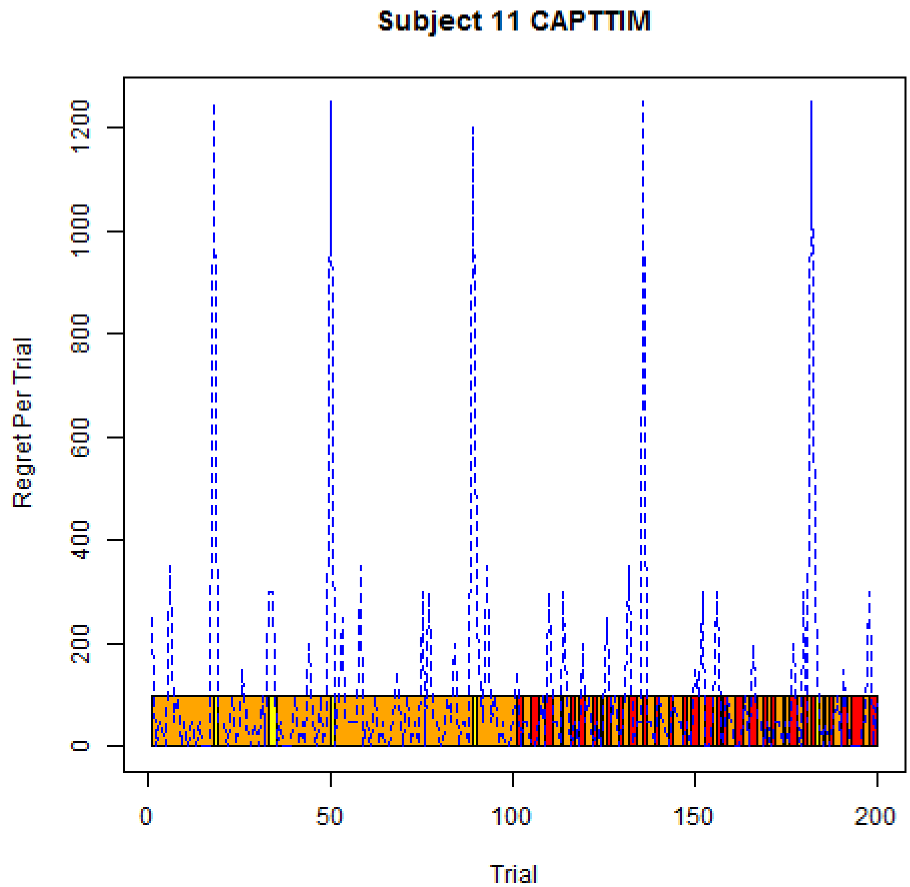


Figure 19. CAPTTIM categorization chart for Subject 11. Note that the values are coded yellow, orange and red. The only reason that Subject 11 was ever categorized as red (high regret and exploitation) within CAPTTIM was due to the fact that subjects are penalized for choosing routes 1 and 2 after trial 100. Subject 11's final damage score was 2200.

Based on the analysis conducted by the research team, the change point analysis of regret provided an accurate delineation between high and low regret. The combination of cognitive state data with the change point analysis in order to generate the CAPTTIM categorization chart is believed to be an effective instructor intervention tool.

C. VALIDATION OF CHANGE POINT ANALYSIS AND COGNITIVE DATA AS CAPTTIM CATEGORIZATION METRICS

All that remained for the research team was to develop a means to validate the effectiveness of using the change point analysis, cognitive state data, and route choice after trial 100. The validation method chosen to validate how well these methods actually categorized subjects within CAPTTIM was a correlation test between number of trials a subject was in the red category and their advantageous selection bias and final damage score. Figures 20 and 21 show the plots for these correlation tests.

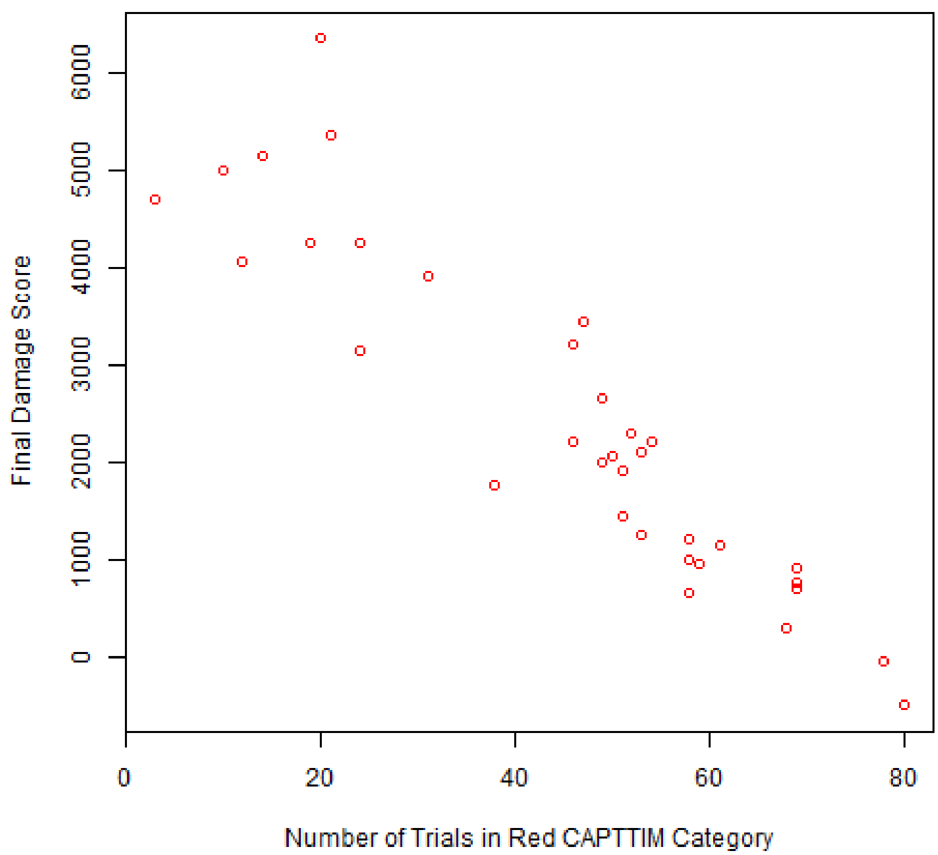


Figure 20. Correlation between final damage score and number of trials spent in the red category of CAPTTIM. The red dots show a strong negative correlation between number of trials spent in the red category and final damage score.

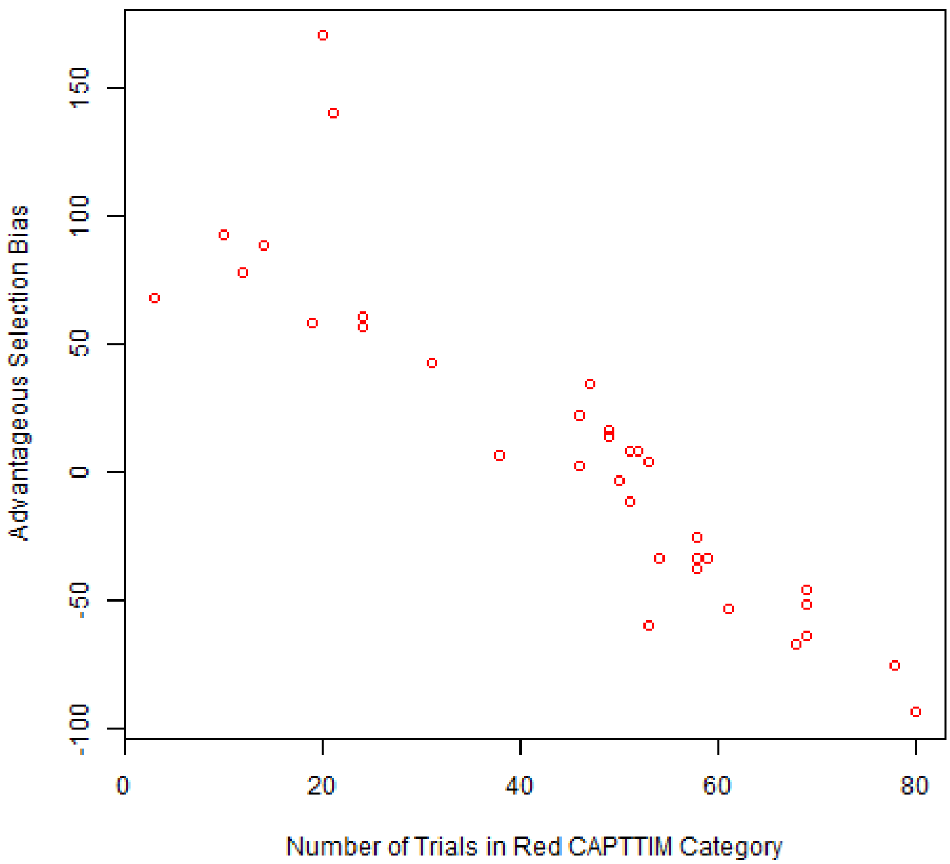


Figure 21. Correlation between advantageous selection bias and number of trials spent in the red category of CAPTTIM. The red dots show a strong negative correlation between number of trials spent in the red category of CAPTTIM and the subject's advantageous selection bias.

The Pearson correlation tests showed a strong negative correlation between the number of trials spent in the red category of CAPTTIM and a subject's final damage score and advantageous selection bias. The correlation test between final damage score and number of trials spent in the red category of CAPTTIM returned a correlation value of – 0.92, $p < .0001$ (95% CI: -0.96 to -0.85), which rejects the null hypothesis that true correlation is equal to 0. The correlation test between advantageous selection bias and number of trials spent

in the red category of CAPTTIM returned a correlation value of – 0.90, p < .0001 (95% CI: -0.95 to -0.81), which rejects the null hypothesis that true correlation is equal to 0.

An additional correlation test was suggested by Dr. Kennedy. Because the number of trials spent in the red and green category of CAPTTIM are not necessarily complementary, the same correlation tests described above were conducted looking at the number of trials spent in the green category of CAPTTIM. Figures 22 and 23 show the plots for these correlation tests.

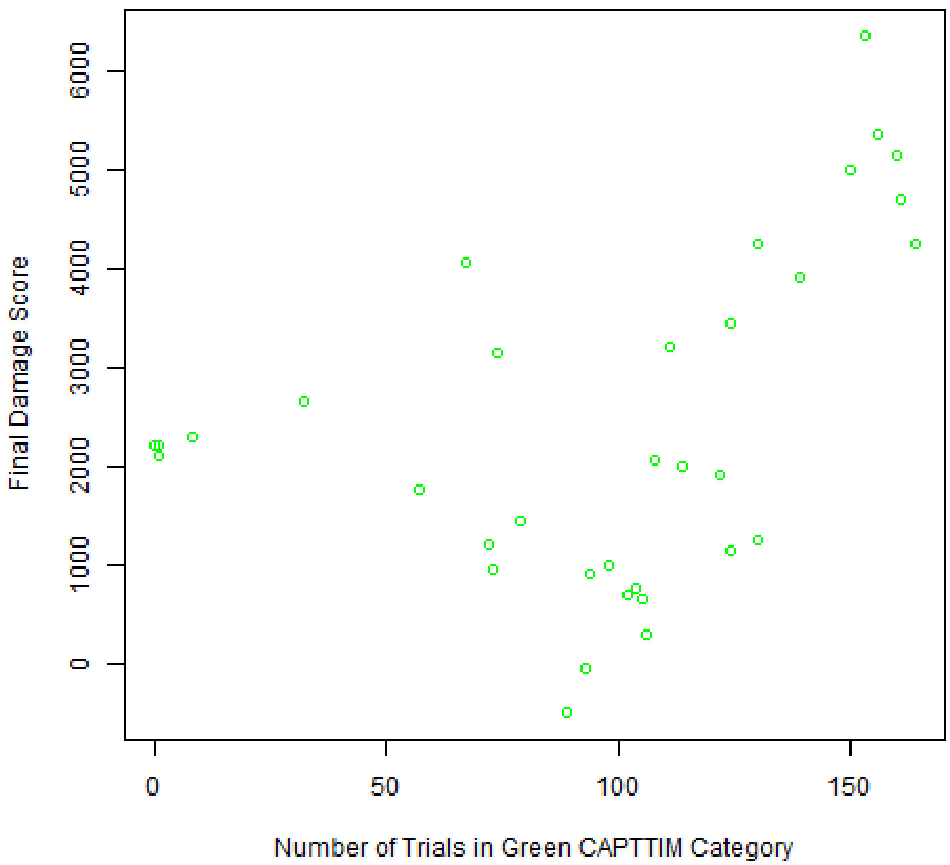


Figure 22. Correlation between final damage score and number of trials spent in the green category of CAPTTIM. The green dots show a moderately strong positive correlation between number of trials spent in the green category and final damage score.

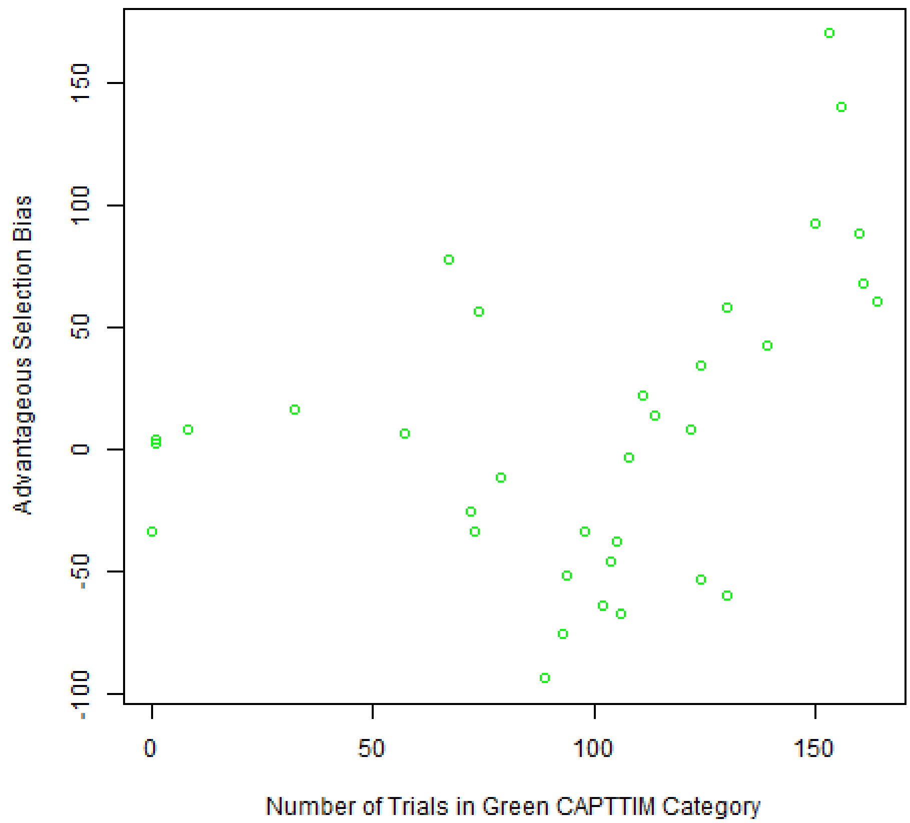


Figure 23. Correlation between advantageous selection bias and number of trials spent in the green category of CAPTTIM. The green dots show a moderately strong positive correlation between number of trials spent in the green category and advantageous selection bias.

Because the plots for these correlations were nonlinear, a Spearman's correlation test was utilized. These tests showed a moderately strong positive correlation between the number of trials spent in the green category of CAPTTIM and a subject's final damage score and advantageous selection bias. The correlation test between final damage score and number of trials spent in the green category of CAPTTIM returned a correlation value of 0.43, $p = .01$, which rejects the null hypothesis that true correlation is equal to 0. The correlation test between advantageous selection bias and number of trials spent in the green

category of CAPTTIM returned a correlation value of 0.38, $p = 0.01$, which rejects the null hypothesis that true correlation is equal to 0.

The weaker correlation between the number of trials spent in the green category of CAPTTIM and final damage score and advantageous selection bias was initially concerning to the research team. However, after further discussion and analysis the weaker correlation made sense. Because the population of high performers (high final damage scores and advantageous selection biases) was smaller within the subject population, the number of trials spent in the green category of CAPTTIM were not as abundant as the number of trials spent in the red category. Additionally, as discussed in the sections above, the third category of subjects were those who never transitioned between the cognitive state of exploration and exploitation. This category of subjects never had the opportunity to experience trials in the green category of CAPTTIM, based on the CAPTTIM categorization algorithm. These observations explained the weaker positive correlation between the numbers of trials spent in the green category compared to the strong negative correlation observed between the numbers of trials spent in the red category.

These results confirmed the use of change point analysis and route choice after trial 100 as an effective method of delineating between high and low regret. When combined with a subject's cognitive state data, these metrics provided an accurate means by which a subject's decision-making pattern could be categorized within the CAPTTIM model.

IV. DISCUSSION

The four primary goals of this thesis were to (1) find a threshold that delineated between high and low regret (decision performance), (2) combine the decision performance data with the cognitive state data, (3) validate these results and CAPTTIM, and (4) develop a visualization method for displaying a subject's CAPTTIM category on a trial by trial basis. All of these primary goals were achieved. This final chapter will summarize the methods used to complete the four primary thesis goals, discuss the implications of the research conducted, discuss future work that could be done to better the CAPTTIM algorithm, and conclude this thesis.

A. SUMMARY OF METHODS USED TO COMPLETE THESIS GOALS

After exploring several analytical approaches, an appropriate method for determining the threshold for regret was found by conducting a change point analysis on the regret per trial that a subject received. The resulting 15 process means returned by the change point analysis were then compared with the median of the subject's 15 process means. The median became the threshold that delineated between high and low regret and categorized the subject's decision performance. An additional metric was introduced based on the number of trials that it took good performers to converge on the ideal decision. On average, the subjects who performed well during the experiment determined that Routes 3 and 4 were the optimal choices by trial 100. Therefore, the additional metric automatically categorized subjects as having high regret if they chose Routes 1 or 2 after trial 100.

This decision performance data was then combined with the cognitive state data that categorized a subject's cognitive state as either exploration or exploitation. The four resulting combinations were (1) high regret and exploration, (2) low regret and exploration, (3) high regret and exploitation, and (4) low regret

and exploitation. As a result of these combinations, a subject's CAPTTIM category could be determined on a trial by trial basis.

The validation of the effectiveness of this CAPTTIM categorization was conducted by performing a Pearson's correlation between the number of trials spent in the red category of CAPTTIM, final damage score, and advantageous selection bias. The Pearson's correlation test was chosen due to the linearity this data exhibited. These correlation results exhibited a very strong negative correlation between these factors. As a result, the number of trials spent in the red category of CAPTTIM proved to be a strong inverse predictor of a subject's final damage score and advantageous selection bias. A Spearman's correlation test was conducted between the number of trials spent in the green category of CAPTTIM, final damage score, and advantageous selection bias. The Spearman's correlation test was chosen due to the nonlinearity this data exhibited. These correlation results showed a moderately strong positive correlation between these factors. As a result, the number of trials spent in the green category of CAPTTIM proved to be a moderate predictor of final damage score and advantageous selection bias.

Finally the visualization of the CAPTTIM category data was designed by creating a bar that exhibited the CAPTTIM category color for each trial. The yellow region of trials is where the subject is experiencing high regret, while their cognitive state is exploration. During a subject's exploration phase, high regret is acceptable and even encouraged. The subject needs to experience high regret in order to gain enough information about the environment to converge and exploit the optimal decision. The orange region of trials is where the subject is experiencing low regret, while their cognitive state is exploration. Long periods of low regret during exploration would require instructor intervention because the subject is ignorantly making the correct decision. Instructor intervention for the orange region could entail letting the subject know that they are making the correct decision or prompting them to sample more of the options to understand

why their decisions are better than the other options. The red region of trials is where a subject is experiencing high regret, while his or her cognitive state is exploitation. Instructor intervention would be required because the subject is exploiting the non-optimal decision believing it to be the optimal decision. The green region of trials is the ideal state in which the subject is experiencing low regret while their cognitive state is exploitation. This yellow, orange, red, and green bar was then overlaid on the regret per trial graph for each subject. This visualization proved to be an effective means of communicating when and where a subject's performance and cognitive state were aligned or misaligned.

B. IMPLICATIONS

The implications of this research are many. CAPTTIM provides feedback on a subject's deviations from the ideal decision path/optimal decision pattern. Based on these deviations, CAPTTIM could provide meaningful feedback to an instructor on the timing and type of intervention that is needed by the trainee. While CAPTTIM is most suited for tasks in which the ideal decision path is known, it could be extrapolated to fit other types of tasks, like rapid response decisions or interactive tactical decision-making games, where understanding optimal decision-making would be beneficial. Another example that CAPTTIM could be extrapolated to fit is wargaming. In wargaming, a commander makes decisions based on the intelligence he/she has received and through trial and error determines the best course of action to execute. The optimal decision path is much more difficult to determine in these examples, but could be determined based on military tactics specific to the wargaming scenario. In these examples inexperienced commanders could conduct wargaming to gain experience that does not involve human lives and receive feedback via CAPTTIM on when and where their performance was aligned or misaligned with their cognitive state.

Another implication of this research is that Army has begun a renewed focus on enhancing the leadership and knowledge of its personnel. The fact that technology has advanced to the degree that countries that used to be inferior in

their military capabilities can now develop quick and innovative solutions that have near peer capabilities, has led the Army to the conclusion that its human resources are its most valuable, adaptable, and flexible assets (Odierno & McHugh, 2015). Based on this conclusion the focus on leadership development tools that train military personnel to be agile, adaptive, and innovative problem solvers in an ambiguous and complex environment has been initiated at the highest level within the Army (Odierno & McHugh, 2015). These leadership development tools range from tasks that aim to improve working memory, comprehending languages, calculating, reasoning, problem solving, and decision-making (Odierno & McHugh, 2015). The ultimate goal of these leadership development tools is to provide technology developed instruction that employs adaptive learning strategies and intelligent tutoring to accelerate learning and education for Soldiers and Army Civilians (Odierno & McHugh, 2015).

The convoy task that was used to collect the data analyzed in this thesis elicits many of the Army's desired leadership development qualities. It requires the user to be adaptive, agile, conduct reasoning, problem solve, and increases working memory and decision-making capabilities. Additionally, the work done in this thesis, specifically the advancement of the model CAPTTIM, has many implications across these leadership development tools. CAPTTIM could be utilized to provide the aspect of intelligent tutoring that could be applied to these technology developed instruction applications that are desired by the Army. Because of CAPTTIM's ability to identify decision performance and cognitive misalignment, it could be used as an intelligent tutor to provide useful feedback to the trainee. Based on these implications the research team believes that CAPTTIM provides a valuable capability to the Army's research on how to develop better leaders and decision makers.

C. FUTURE WORK

As previously stated the delineation between high and low regret and the cognitive states of exploration and exploitation was calculated retrospectively. In order for CAPTTIM to be able to provide “real-time” feedback to an instructor or even a trainee, these delineations must be calculated dynamically. This is the most crucial advancement that must take place in this research in order for CAPTTIM to be a more effective tool for instructors. One way that this can be accomplished is to have a “burn in period” that is a set number of trials where no feedback is provided and a subject is not categorized into any CAPTTIM category. Once this period is complete, a change point analysis of regret per trial can be performed to determine the threshold that delineates between high and low regret. After this threshold is calculated for this period, all future decision performance can be compared to that threshold on a trial by trial basis. The same concept applies to the EWMA of latency in decision-making times in order to provide the delineation between the cognitive states of exploration and exploitation. Once this threshold is calculated for the “burn in period” a subject can be categorized into one of the two cognitive states on subsequent trials. These two delineations can then be combined, as they were in this thesis, to categorize subjects into a CAPTTIM category. An initial analysis of this “burn in period” concept with the research team, suggested that a period of 50–80 trials would be sufficient to calculate a threshold for decision performance and cognitive states.

Other future work would be to (1) test CAPTTIM on a task that differs from the convoy task, and (2) develop the CAPTTIM oriented intervention feedback loop. Testing CAPTTIM on a task like wargaming, rapid response decisions, or tactical decision-making games will help validate CAPTTIM’s adaptability to different tasks. By validating the adaptability of CAPTTIM, the significance of this research to the Army’s leadership development focus will be further solidified. The development of the CAPTTIM oriented intervention feedback loop is necessary to enable the model to be used as an intelligent tutor in computer

based tasks. The ability for a script to be created that utilizes data categorized by CAPTTIM and provides task specific guidance/feedback to a trainee will, again, further illustrate CAPTTIM's implication to the Army's leadership development program.

D. CONCLUSION

Understanding optimal decision-making is a very difficult task, but one that is worth undertaking. The Army and the military as a whole have realized that, due to budget constraints, they are entering into one of the most fiscally austere environments that the military has experienced in decades (Odierno & McHugh, 2015). As a result, they have grasped that the dominance of the United States military will not be accomplished by the unlimited acquisition of newer weapons, vehicles, and technology (Odierno & McHugh, 2015). Thus military dominance will be measured by the ability to develop military professionals that are capable of being effective, agile, adaptive, and innovative decision makers and problem solvers (Odierno & McHugh, 2015). The focus on force development versus the acquisition of material solutions lends gravity to the research conducted in this thesis.

The research team believes that the work done in this thesis has furthered the understanding of decision-making and directly provides a useful tool that could be used to aid leadership development programs. While there is still much to discover when it comes to understanding how humans process information and make decisions, this research has made it more possible to understand and classify decision performance and cognitive state. With this understanding the human mind becomes less of a black box, in which an instructor or intelligent tutor has no insight, and allows a small peek at what is really going on in the subject's decision-making process. This peek is made possible by the ability to understand and identify the alignment or misalignment of cognitive state with decision performance. By looking at a common reinforcement learning task, modified for the military domain, the research team was able to investigate and

better understand a subject’s decision-making pattern and how to intelligently influence this pattern if determined to be suboptimal. It will be exciting to see what follow on research discovers, and how CAPTTIM is modified to increase the understanding of optimal decision-making.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX A. PAYOUT SCHEDULE FOR IGT AND CONVOY TASK

IGT Payout Schedule			
Deck A	Deck B	Deck C	Deck D
-150	100	50	50
-250	100	0	50
100	100	50	50
100	100	0	50
-50	100	50	50
100	100	0	50
-200	100	50	50
100	100	0	50
-100	-1150	0	50
100	100	0	-200
-150	100	50	50
-250	100	50	50
100	100	0	50
-50	100	50	50
100	100	0	50
-200	100	50	50
100	100	0	50
-100	100	50	50
100	-1150	0	50
-150	100	0	-200
-250	100	50	50
100	100	50	50
100	100	0	50
-50	100	50	50
100	100	0	50
-200	100	50	50
100	100	0	50
-100	100	50	50
100	-1150	0	50
-150	100	0	-200
-250	100	50	50
100	100	50	50
100	100	0	50
-50	100	50	50

Convoy Task Payout Schedule			
Rout 1	Route 2	Route 3	Route 4
-150	100	50	50
-250	100	0	50
100	100	50	50
100	100	0	50
-50	100	50	50
100	100	0	50
-200	100	50	50
100	100	0	50
-100	-1150	0	50
100	100	0	-200
-150	100	50	50
-250	100	50	50
100	100	0	50
-50	100	50	50
100	100	0	50
-200	100	50	50
100	100	0	50
-100	100	50	50
100	-1150	0	50
-150	100	0	-200
-250	100	50	50
100	100	50	50
100	100	0	50
-50	100	50	50
100	100	0	50
-200	100	50	50
100	100	0	50
-100	100	50	50
100	-1150	0	50
-150	100	0	-200
-250	100	50	50
100	100	50	50
100	100	0	50
-50	100	50	50

100	100	0	50
-200	100	50	50
100	100	0	50
-100	100	50	50
100	-1150	0	50
-150	100	0	-200
-250	100	50	50
100	100	50	50
100	100	0	50
-50	100	50	50
100	100	0	50
-200	100	50	50
100	100	0	50
-100	100	50	50
100	-1150	0	50
-150	100	0	-200
-250	100	50	50
100	100	50	50
100	100	0	50
-50	100	50	50
100	100	0	50
-200	100	50	50
100	100	0	50
-100	100	50	50
100	-1150	0	50
-150	100	0	-200
-250	100	50	50
100	100	50	50
100	100	0	50
-50	100	50	50
100	100	0	50
-200	100	50	50
100	100	0	50
-100	100	50	50
100	-1150	0	50
-150	100	0	-200
-250	100	50	50
100	100	50	50
100	100	0	50
-50	100	50	50

100	100	0	50
-200	100	50	50
100	100	0	50
-100	100	50	50
100	-1150	0	50
-150	100	0	-200
-250	100	50	50
100	100	50	50
100	100	0	50
-50	100	50	50
100	100	0	50
-200	100	50	50
100	100	0	50
-100	100	50	50
100	-1150	0	50
-150	100	0	-200
-250	100	50	50
100	100	50	50
100	100	0	50
-50	100	50	50
100	100	0	50
-200	100	50	50
100	100	0	50
-100	100	50	50
100	-1150	0	50
-150	100	0	-200
-250	100	50	50
100	100	50	50
100	100	0	50
-50	100	50	50
100	100	0	50
-200	100	50	50
100	100	0	50
-100	100	50	50
100	-1150	0	50
-150	100	0	-200
-250	100	50	50
100	100	50	50
100	100	0	50
-50	100	50	50

100	100	0	50
-200	100	50	50
100	100	0	50
-100	100	50	50
100	-1150	0	50
-150	100	0	-200
-250	100	50	50
100	100	50	50
100	100	0	50
-50	100	50	50
100	100	0	50
-200	100	50	50
100	100	0	50
-100	100	50	50
100	-1150	0	50
-150	100	0	-200
-250	100	50	50
100	100	50	50
100	100	0	50
-50	100	50	50
100	100	0	50
-200	100	50	50
100	100	0	50
-100	100	50	50
100	-1150	0	50

100	100	0	50
-200	100	50	50
100	100	0	50
-100	100	50	50
100	-1150	0	50
-150	100	0	-200
-250	100	50	50
100	100	50	50
100	100	0	50
-50	100	50	50
100	100	0	50
-200	100	50	50
100	100	0	50
-100	100	50	50
100	-1150	0	50
-150	100	0	-200
-250	100	50	50
100	100	50	50
100	100	0	50
-50	100	50	50
100	100	0	50
-200	100	50	50
100	100	0	50
-100	100	50	50
100	-1150	0	50
-150	100	0	-200
-250	100	50	50
100	100	50	50
100	100	0	50
-50	100	50	50
100	100	0	50
-200	100	50	50
100	100	0	50
-100	100	50	50
100	-1150	0	50
-150	100	0	-200
-250	100	50	50
100	100	50	50
100	100	0	50
-50	100	50	50

100	100	0	50
-200	100	50	50
100	100	0	50
-100	100	50	50
100	-1150	0	50
-150	100	0	-200
-250	100	50	50
100	100	50	50
100	100	0	50
-50	100	50	50
100	100	0	50
-200	100	50	50
100	100	0	50
-100	100	50	50
100	-1150	0	50
-150	100	0	-200
-250	100	50	50
100	100	50	50
100	100	0	50
-50	100	50	50
100	100	0	50
-200	100	50	50
100	100	0	50
-100	100	50	50
100	-1150	0	50
-150	100	0	-200
-250	100	50	50
100	100	50	50
100	100	0	50
-50	100	50	50
100	100	0	50
-200	100	50	50
100	100	0	50
-100	100	50	50
100	-1150	0	50
-150	100	0	-200
-250	100	50	50
100	100	50	50
-50	100	0	50
100	100	50	50

-200	100	0	50
100	100	50	50
-100	100	0	50
100	100	50	50
-150	-1150	0	50
-250	100	0	-200
100	100	50	50
100	100	50	50
-50	100	0	50
100	100	50	50
-200	100	0	50
100	100	50	50
-100	100	0	50
100	100	50	50
-150	-1150	0	50
-250	100	0	-200
100	100	50	50
100	100	50	50
-50	100	0	50
100	100	50	50
-200	100	0	50
100	100	50	50
-100	100	0	50
100	100	50	50
-150	-1150	0	50
-250	100	0	-200
100	100	50	50
100	100	50	50
-50	100	0	50
100	100	50	50
-200	100	0	50
100	100	50	50
-100	100	0	50
100	100	50	50
-150	-1150	0	50
-250	100	0	-200
100	100	50	50
100	100	50	50
-50	100	0	50
100	100	50	50

-200	100	0	50
100	100	50	50
-100	100	0	50
100	100	50	50
-150	-1150	0	50
-250	100	0	-200
100	100	50	50
100	100	50	50
-50	100	0	50
100	100	50	50
-200	100	0	50
100	100	50	50
-100	100	0	50
100	100	50	50
-150	-1150	0	50
-250	100	0	-200
100	100	50	50
100	100	50	50
-50	100	0	50
100	100	50	50
-200	100	0	50
100	100	50	50
-100	100	0	50
100	100	50	50
-150	-1150	0	50
-250	100	0	-200
100	100	50	50
100	100	50	50
-50	100	0	50
100	100	50	50
-200	100	0	50
100	100	50	50
-100	100	0	50
100	100	50	50
-150	-1150	0	50
-250	100	0	-200
100	100	50	50
100	100	50	50
-50	100	0	50
100	100	50	50

-200	100	0	50
100	100	50	50
-100	100	0	50
100	100	50	50
-150	-1150	0	50
-250	100	0	-200
100	100	50	50
100	100	50	50
-50	100	0	50
100	100	50	50
-200	100	0	50
100	100	50	50
-100	100	0	50
100	100	50	50
-150	-1150	0	50
-250	100	0	-200
100	100	50	50
100	100	50	50
-50	100	0	50
100	100	50	50
-200	100	0	50
100	100	50	50
-100	100	0	50
100	100	50	50
-150	-1150	0	50
-250	100	0	-200
100	100	50	50
100	100	50	50
-50	100	0	50
100	100	50	50
-200	100	0	50
100	100	50	50
-100	100	0	50
100	100	50	50
-150	-1150	0	50
-250	100	0	-200
100	100	50	50
100	100	50	50
-50	100	0	50
100	100	50	50

-200	100	0	50
100	100	50	50
-100	100	0	50
100	100	50	50
-150	-1150	0	50
-250	100	0	-200
100	100	50	50
100	100	50	50
-50	100	0	50
100	100	50	50
-200	100	0	50
100	100	50	50
-100	100	0	50
100	100	50	50
-150	-1150	0	50
-250	100	0	-200
100	100	50	50
100	100	50	50
-50	100	0	50
100	100	50	50
-200	100	0	50
100	100	50	50
-100	100	0	50
100	100	50	50
-150	-1150	0	50
-250	100	0	-200

APPENDIX B. R SCRIPTS

A. EWMA OF DECISION LATENCY TIMES R SCRIPT

```
print("begin script: ODM multi-arm bandit analysis")
setwd("~/NPS/Thesis/Thesis Data/Data Critz")
require(zoo)
require(ggplot2)
require(fTrading)
require(qcc)
require(RColorBrewer)
require(StatMatch)
IGT <- T # Are we using the published IGT payout schedule?
PlayerInput <- T # Are we analysing a human player?
doRegretA.mb <- T # regret by absolute

Basics <- F # plot basic histograms
BasicsT <- F # plot basic histograms

# Create, test through MC, plot new distributions...

numTrials <-200 #  ignore any more than 200 trials
cog.state <- vector() #Capture cognitive state data
route.select <- vector() #Capture route choice

# Read in payout schedule
  IGTresponse <- read.csv("IGTimproved.csv")
  numBandits = length(IGTresponse)
  numTrials <-200

# Read in player input
if (PlayerInput){
  files <- list.files(pattern = '*MultiArmBandit*')
  numPlayers <- length(files)
  numBandits <- 4
  subject <- 1
  # Create dataframe for subject specific response
  MA.decision <- data.frame(matrix(0,nrow=200,ncol=numPlayers))
  # Create dataframe for descriptive statistics
  MA.summary <- data.frame(matrix(0,nrow=numPlayers,ncol=35))
  header <-c('Subject','mb.FD.100','mb.numFD.100','mb.numHFD.100',
           'mb.R1.100','mb.R2.100','mb.R3.100','mb.R4.100','mb.adv.sb.100',
```

```
'mb.mean.l.100','mb.med.l.100','mb.sd.100','mb.numFD.SecHalf','mb.numHFD.S
ecHalf',

'mb.R1.SecHalf','mb.R2.SecHalf','mb.R3.SecHalf','mb.R4.SecHalf','mb.adv.sb.Se
cHalf',

'mb.mean.l.SecHalf','mb.med.l.SecHalf','mb.sd.SecHalf','mb.FD.200','mb.numFD.
200','mb.numHFD.200',
          'mb.R1.200','mb.R2.200','mb.R3.200','mb.R4.200','mb.adv.sb.200',
          'mb.mean.l.200','mb.med.l.200','mb.sd.200','SigLat','perc.regret')
  names(MA.summary) <- header

  # df used for calculating regret
  Regret.mb.df <- data.frame(matrix(0,nrow=0,ncol=5))


  #Import Player choices and resulting response by trial
  #file <- files[1]
  element<-1
  for(file in files){

    PlayerID <- file#paste('Subject ',subject,sep="")
    print(PlayerID)
    player <- read.csv(file)
    #print(summary(player))
    LL <- list()
    player<- subset(player, trial<201)
    numTrials <- length(player[,1])

    # add players decision to MA.decision
    colnames(MA.decision)[element]<-as.numeric(noquote(strsplit(PlayerID,"
")[[1]])[1])
    MA.decision[element] <- player$routeSel
    decide <- as.numeric(player$routeSel)      # get decision data)
    decide[decide== "1"] <- -1 # recode selections to adv sel scores
    decide[decide== "2"] <- -1
    decide[decide== "3"] <-  1
    decide[decide== "4"] <-  1
    element<-element+1

    # Latency by trial number plot
    numShift      <-numTrials-1
    shift         <-append(0,head(player$trialLoss,numShift),after=1)
    Damage.before <-factor(player$trialLoss)
```

```
   Damage.after  <-factor(shift)
   size.before   <-factor(player$trialLoss)
   size.after    <-factor(shift)
   Damage.color  <-factor(player$trialLoss)
   damage.cat       <-list('none to low (0,50)'=0,'none to low (0,50)'=50,'med
(150,200,250,300,350)'=150,'med (150,200,250,300,350)'=200,
               'med                     (150,200,250,300,350)'=250,'med
(150,200,250,300,350)'=300,'med           (150,200,250,300,350)'=350,'high
(1250)'=1250)
   damage.size<-
list('10'=0,'10'=50,'20'=150,'20'=200,'20'=250,'20'=300,'20'=350,'100'=1250)
   damage.color<-
list('3'=0,'3'=50,'2'=150,'2'=200,'2'=250,'2'=300,'2'=350,'5'=1250)
   levels(Damage.before) <- damage.cat
   levels(Damage.after)  <- damage.cat
   levels(size.before)   <- damage.size
   levels(size.after)    <- damage.size
   levels(Damage.color)  <- damage.color
   myColors             <- brewer.pal(5,"Set1")
   names(myColors)       <- c(100,20,10)
   colScale <- scale_colour_manual(name = "damage",values = myColors)


   player<-
cbind(player,Damage.before,Damage.after,size.before,size.after)#,ewmaS)

###Fill in summary stats for 100 trials
   #'Subject'
   subject <- as.numeric(noquote(strsplit(PlayerID, " ")[[1]])[1])
   MA.summary[subject,1]<- subject
   #'Final Damage'
   MA.summary[subject,2]<- player$Damage[100]
   #'# trials friendly damage'
   MA.summary[subject,3]<- sum(player$trialLoss[1:100]>0)
   #'# trials heavy friendly damage'
   MA.summary[subject,4]<- sum(player$trialLoss[1:100]>1000)
   #'Route 1'
   MA.summary[subject,5]<- sum(player$routeSel[1:100]=='1')/100
   #'Route 2'
   MA.summary[subject,6]<- sum(player$routeSel[1:100]=='2')/100
   #'Route 3'
   MA.summary[subject,7]<- sum(player$routeSel[1:100]=='3')/100
   #'Route 4'
   MA.summary[subject,8]<- sum(player$routeSel[1:100]=='4')/100
   #'advantageuos selection bias'
```

```
   MA.summary[subject,9]<-
sum(player$routeSel[1:100]=='3')+sum(player$routeSel[1:100]=='4')-
sum(player$routeSel[1:100]=='1')-sum(player$routeSel[1:100]=='2')
   #'mean latency time'
   MA.summary[subject,10]<- mean(player$latent[2:100])
   #'median latency'
   MA.summary[subject,11]<- median(player$latent[2:100])
   #'standard deviation latency'
   MA.summary[subject,12]<- sd(player$latent[2:100])

   #Fill in summary stats for second half, 101-200 trials
   #'# trials friendly damage'
   MA.summary[subject,13]<- sum(player$trialLoss[101:200]>0)
   #'# trials heavy friendly damage'
   MA.summary[subject,14]<- sum(player$trialLoss[101:200]>1000)
   #'Route 1'
   MA.summary[subject,15]<- sum(player$routeSel[101:200]=='1')/100
   #'Route 2'
   MA.summary[subject,16]<- sum(player$routeSel[101:200]=='2')/100
   #'Route 3'
   MA.summary[subject,17]<- sum(player$routeSel[101:200]=='3')/100
   #'Route 4'
   MA.summary[subject,18]<- sum(player$routeSel[101:200]=='4')/100
   #'advantageuos selection bias'
   MA.summary[subject,19]<-
sum(player$routeSel[101:200]=='3')+sum(player$routeSel[101:200]=='4')-
sum(player$routeSel[101:200]=='1')-sum(player$routeSel[101:200]=='2')
   #'mean latency time'
   MA.summary[subject,20]<- mean(player$latent[101:200])
   #'median latency'
   MA.summary[subject,21]<- median(player$latent[101:200])
   #'standard deviation latency'
   MA.summary[subject,22]<- sd(player$latent[101:200])

   #Fill in summary stats for 200 trials
   #'Final Damage'
   MA.summary[subject,23]<- player$Damage[numTrials]
   #'# trials friendly damage'
   MA.summary[subject,24]<- sum(player$trialLoss>0)
   #'# trials heavy friendly damage'
   MA.summary[subject,25]<- sum(player$trialLoss>1000)
   #'Route 1'
   MA.summary[subject,26]<- sum(player$routeSel=='1')/numTrials
   #'Route 2'
   MA.summary[subject,27]<- sum(player$routeSel=='2')/numTrials
```

```
#'Route 3'
MA.summary[subject,28]<- sum(player$routeSel=='3')/numTrials
#'Route 4'
MA.summary[subject,29]<- sum(player$routeSel=='4')/numTrials
#'advantageuos selection bias'
MA.summary[subject,30]<-
sum(player$routeSel=='3')+sum(player$routeSel=='4')-
sum(player$routeSel=='1')-sum(player$routeSel=='2')
#'mean latency time'
MA.summary[subject,31]<- mean(player$latent[2:200])
#'median latency'
MA.summary[subject,32]<- median(player$latent[2:200])
#'standard deviation latency'
MA.summary[subject,33]<- sd(player$latent[2:200])
#'Significant latency'
MA.summary[subject,34]<- mean(player$latent[player$size.before==100])

if(doRegretA.mb){
  num.a <- 1 # set the next trial to one for each option
  num.b <- 1
  num.c <- 1
  num.d <- 1
  regret.total <- 0 # initialize total regret
  regret.c <- 0 # initialize regret count
  regret.r <- 0 # initialize regret rate
  for(trial in 1:numTrials){ # for every trial (withing every player loop)
    # The best option value (gain+loss already computed) in the schedule for
each option
    opt.choice.v<-
max(IGTresponse[num.a,1],IGTresponse[num.b,2],IGTresponse[num.c,3],IGTres
ponse[num.d,4])
    # From the records, what they gained and lost
    player.choice.v <- player$trialGain[trial]-player$trialLoss[trial] # positive is
good
    # find the difference
    regret.v <- opt.choice.v - player.choice.v
    if(regret.v>0){regret.c <- regret.c +1}
    regret.r <- regret.c/trial
    # accumulate regret
    regret.total <- regret.total + regret.v
    # normalize by trials
    regret.mean <- regret.total / trial
    # error check
```

```
       #
if(regret.v<0){print(paste(num.a,num.b,num.c,num.d,'opt',opt.choice.v,'player',pla
yer.choice.v,'regret =',regret.v,' sub ',subject,' trial ',trial))}
       # update next available options
       if(player$routeSel[trial]==1){num.a<-num.a+1}
       if(player$routeSel[trial]==2){num.b<-num.b+1}
       if(player$routeSel[trial]==3){num.c<-num.c+1}
       if(player$routeSel[trial]==4){num.d<-num.d+1}
       # combine into row
       trial.regret<-
c(trial,decide[trial],regret.v,regret.total,regret.mean,subject,regret.r)
       # add to Regret.df data.frame of all trial/regret measure/player combinations
       Regret.mb.df <- rbind(Regret.mb.df,trial.regret)
     }
   }
#'Significant latency'
MA.summary[subject,35]<- regret.r

player <- player[-1,] # Remove first lantency observation

### Sequential Detection Methods for Detecting Exploration-Exploitation Mode
Changes

  ### Method 1: The Exponentially Weighted Moving Average

# develop single number of standard deviation of all latencies after low damage
threshold <- 2 # threshold multiplier
mb.sd.threshold <- sd(player$latent[player$size.before==10])*threshold

# develop estimate of moving latency from exponential moving z_t = ?? y_t + (1-
??) z_{t-1}
EWMAlambda <- .1 # lambda
ewma.latent.lst<-
ewmaSmooth(player$trial[player$size.before==10],player$latent[player$size.befo
re==10],lambda=EWMAlambda) # list of estimate data

# build a dataframe with this data in it
EWMA <- data.frame(matrix('NA',nrow=length(ewma.latent.lst$x),ncol=3))
header <-c('trial','ewma','threshold')
names(EWMA) <- header
EWMA['trial'] <- ewma.latent.lst$x
EWMA['ewma'] <- ewma.latent.lst$y
EWMA['threshold'] <- mb.sd.threshold

# merge it with the other player data
```

```
player <- merge(player,EWMA,by="trial",all.x=T,fill=NA)

# Inpute data from missing high damage +1 trials
# input by 'hot deck', simply continue last value until next observation (estimate in
this case)
ewma.shift<-append(0,head(player$ewma,length(player$ewma)-1),after=1)
#vector from shifting ewma down 1
num.mistakes <-5
for(mistake in 1:num.mistakes){
ewma.shift<-append(0,head(ewma.shift,length(ewma.shift)-1),after=1)#shift
again...
player$ewma[is.na(player$ewma)]<-ewma.shift[is.na(player$ewma)]
}

# build upper and lower bounds for colored ribbons on graph
player['upper.line'] <- apply(cbind(player$threshold,player$ewma),1,max)
player['lower.line'] <- apply(cbind(player$threshold,player$ewma),1,min)
cog.stateTmp <- numeric(200)
cog.stateTmp[1] <- "explore"
cog.stateTmp[2:200] <- ifelse(player$ewma>player$threshold,"explore","exploit")
cog.state <- c(cog.state,cog.stateTmp)
#Due to long latency, we do not count the first route selection.
route.selectTmp <- numeric(200)
route.selectTmp[1] <- 0 #Can be any value for this analysis
route.selectTmp[2:200] <- player$routeSel
route.select <- c(route.select,route.selectTmp)

### Method 2: Monitoring Sequential Sample Variances

###Create / Save graphs for each subject
#    maxLatent <- 8
#    gtitle <- paste('Latency and EWMA by trial number for',PlayerID)
#    ftitle <- paste0(subject,'TxL.png')
#    LatByTrial<-ggplot(data=player,aes(x=trial,y=latent))+
#
geom_ribbon(aes(ymin=threshold,ymax=upper.line,linetype="NA"),fill="orange",al
pha=.5,show_guide=F)+
#
geom_ribbon(aes(ymin=lower.line,ymax=threshold,linetype="NA"),fill="skyblue",a
lpha=.5,show_guide=F)+
#
labs(title=gtitle)+coord_cartesian(ylim=c(0,maxLatent))+colScale+theme_bw()+xl
ab("Trials")+ylab("Latency")
```

```
#                                                         LatByTrial<-
LatByTrial+geom_line(data=player,aes(x=trial,y=ewma),linetype=1,colour="grey8
8")
#                                                         LatByTrial<-
LatByTrial+geom_point(data=player,aes(x=trial,y=latent,color=size.after,size=siz
e.after),show_guide=T)
#    #png(file=ftitle,width = 1000, height = 700)
#    print(LatByTrial)
#    maxLatent <- 8
#    gtitle <- paste('Latency and EWMA by trial number for',PlayerID)
#    ftitle <- paste0(subject,'TxL.png')
#    LatByTrial<-ggplot(data=player,aes(x=trial,y=latent))+
#
geom_ribbon(aes(ymin=threshold,ymax=upper.line,linetype=NA,fill="Explore"),al
pha=.5,show_guide=T)+
#
geom_ribbon(aes(ymin=lower.line,ymax=threshold,linetype=NA,fill="Exploit"),alp
ha=.5,show_guide=F)+
#    scale_fill_manual(values=c("Explore"='orange',"Exploit"="skyblue"))+
#
#labs(title=gtitle)+coord_cartesian(ylim=c(0,maxLatent))+theme_bw()+xlab("Trial
s")+ylab("Latency")
#
labs(title=gtitle)+coord_cartesian(ylim=c(0,maxLatent))+colScale+theme_bw()+xl
ab("Trials")+ylab("Latency")
#LatByTrial<-
LatByTrial+geom_line(data=player,aes(x=trial,y=ewma),linetype=1,colour="grey8
8")
#LatByTrial<-
LatByTrial+geom_point(data=player,aes(x=trial,y=latent,color=size.after,size=siz
e.after),show_guide=T)
#    #png(file=ftitle,width = 1000, height = 700)
#
#    print(LatByTrial)
#    dev.off()
#
#    gtitle <- paste('Route by trial number for',PlayerID)
#      plotBT<- ggplot(player,aes( trial,colour = size.before,factor(routeSel))) +
labs(title = gtitle)+colScale
#      plotBT<-plotBT+geom_point(aes(size = size.before),show_guide = F) +
theme_bw()+  xlab("Trials") +ylab("Routes")
#              #plotBT<-plotBT+geom_point(aes(colour = Damage.color))#+
scale_fill_continuous(name = "Friendly damage on previous
trial")#+coord_cartesian(ylim=c(0,8))
```

```
#      plotBT<-plotBT + theme(legend.direction = "horizontal", legend.position =
"bottom")#+annotate("text", x = 0, y = 10, label = "Relationship between x and y")
#      #LatByTrial+ guides(fill = guide_legend(title.theme = element_text(size=15,
face="italic", colour = "red", angle = 45)))
#    ftitle <- paste0(subject,'TxR.png')
#    png(file=ftitle,width = 1000, height = 700)
#    suppressWarnings(print(plotBT))
#    dev.off()

   subject <- subject+1

  }

header<-
c('trial','adv.sel.bias','regret.trial','regret.total','regret.mean','subject','regret.rate')
names(Regret.mb.df) <- header
  } # end of read in player input (PlayerInput)

survey_data<-
merge(read.csv("survey_data.csv"),read.csv("groups.csv"),by="Subject")
total<-merge(survey_data,MA.summary,by="Subject")
Regret.mb.df$Cog.State <- cog.state
Regret.mb.df$RouteSel <- route.select

save.image("C:/Users/John/Documents/NPS/Thesis/ThesisData/Data
Critz/RegretData.RData")
```

B. CHANGEPOINT ANALYSIS R SCRIPT

```
setwd("~/NPS/Thesis/Thesis Data/Data Critz")
load("C:/Users/John/Documents/NPS/Thesis/ThesisData/Data
Critz/RegretData.RData")

library("changepoint")
subject.vec <- unique(Regret.mb.df$subject) #For all subjects
#subject.vec <- subject.vec[9]
#subject.vec <- c(1,4,8,11,14,15,17,26,28)
regret.vec <- numeric(200)
median.vec <- numeric (200)
med.dev <- numeric(200)
#upperCTLLimit <- numeric(200)
bin <- list()
chngepoint.bin <- list()
bin.vec <- numeric(200)
subject.index <- 1
subject.start <- 1
```

```
subject.difference <- 200

for(index in 1:length(subject.vec)){
  subject.tmp <- which(Regret.mb.df$subject==subject.vec[index])
  test.subj <- Regret.mb.df[subject.tmp[1]:subject.tmp[200],]
#   a <- 1
#   b <- 5
  bin.index <- 1
  tmp.chng <- cpt.mean(test.subj[,3], method="SegNeigh",Q=15)
  chngepoint.bin[[index]] <- tmp.chng
  #Corrected histogram label
  png(paste("RegretHistogramSubject",subject.vec[index],".png",sep=""))
  hist(test.subj[,3],col="blue",xlab="Regret  Value",main=paste("Regret  Histogram
for Subject ",subject.vec[index],sep=""))
  dev.off()
}

save.image("C:/Users/John/Documents/NPS/Thesis/ThesisData/Data
Critz/RegretData.RData")
```

C. CAPTTIM VISUALIZATION R SCRIPT

```
#Had to create the vector for subject 9 manually
#Source Revised MultiArm
#Source Regret.Mean file
require(data.table) #Required to find unique column elements
#Find the sujects we want
#subject.vec <- unique(Regret.mb.df$subject) #For all subjects
#subject.vec <- c(1,4)
#subject.vec <- c(11)
#index <- 1
#subject.vec <- subject.vec[-c(1:8)]
#subject.vec1 <- subject.vec[-9]

subject.control.vec1 <- vector()
subject.category1 <- vector()
index <- 1
for(index in 1:length(subject.vec)){
  print(paste("Processing Subject ",subject.vec[index]))
  subject.tmp <- which(Regret.mb.df$subject==subject.vec[index])
  test <- Regret.mb.df[subject.tmp[1]:subject.tmp[200],]
  test2 <- chngepoint.bin[[index]]
  chgptmean.vec <- numeric(200) #Creat a vector to collect the changepoints
  i <- 1
  while(i < length(test2@cpts)+1){
    #   browser()
```

```
  #    print(paste("I is ",i))
  #    print(chgptmean.vec)
  if(i==1){
    chgptmean.vec[i] <- test2@param.est$mean[i]
    i <- i + 1
    next
  }
  if(test2@cpts[i]!=200){
    if(test2@cpts[i]-test2@cpts[i-1]==1){
      chgptmean.vec[test2@cpts[i]] <- test2@param.est$mean[i]
      i <- i + 1
      next
    }
    if(test2@cpts[i+1]-test2@cpts[i]==1){
      chgptmean.vec[(test2@cpts[i-1]+1):(test2@cpts[i])]<-
test2@param.est$mean[i]
      i <- i + 1
      next
    }

    if(test2@cpts[i+1]-test2@cpts[i]>1){
      chgptmean.vec[(test2@cpts[i-1]+1):(test2@cpts[i])]<-
test2@param.est$mean[i]
      i <- i + 1
      next
    }
  }

  if(test2@cpts[i]==200){
    chgptmean.vec[(test2@cpts[i-1]+1):(test2@cpts[i])]<-
test2@param.est$mean[i]
    i <- i+1
  }

 }

 test$Mean.Regret <- chgptmean.vec #Add this to whatever dataframe you
would like of the same length
 #Now let's add color
 #First let’s find out which trials were in or out of control
 control.vec <- numeric(200)
 for(i in 1:200){
  if(test$Mean.Regret[i]>median(test2@param.est$mean)){
    control.vec[i] <- "high"
  }
```

```
  if(test$Mean.Regret[i]<=median(test2@param.est$mean)) {
    control.vec[i] <- "low"
  }
}

test$Control <- control.vec
subject.control.vec1 <- c(subject.control.vec1,control.vec)

#Next, make up a color for each value
color.vec <- numeric(200)
for(i in 1:200){
  if(i <= 100){
    if(test$Cog.State[i]=='explore' & test$Control[i]=="low"){
      color.vec[i] <- "orange"
    }
    if(test$Cog.State[i]=='explore' & test$Control[i]=="high") {
      color.vec[i] <- "yellow"
    }
    if(test$Cog.State[i]=='exploit' & test$Control[i]=="low") {
      color.vec[i] <- "green"
    }
    if(test$Cog.State[i]=='exploit' & test$Control[i]=="high") {
      color.vec[i] <- "red"
    }
  }
  if(i > 100){
    if(test$RouteSel[i]==2) {
      color.vec[i] <- "red"
      next
    }
    if(test$RouteSel[i]==1) {
      color.vec[i] <- "red"
      next
    }
    if(test$Cog.State[i]=='explore' & test$Control[i]=="low"){
      color.vec[i] <- "orange"
    }
    if(test$Cog.State[i]=='explore' & test$Control[i]=="high") {
      color.vec[i] <- "yellow"
    }
    if(test$Cog.State[i]=='exploit' & test$Control[i]=="low") {
      color.vec[i] <- "green"
    }
    if(test$Cog.State[i]=='exploit' & test$Control[i]=="high") {
      color.vec[i] <- "red"
```

```
    }
   }


  }
  #test$Color <- color.vec
  subject.category1 <- c(subject.category1,color.vec)
  test$Color <- color.vec
  png(paste("Subject",subject.vec[index],"CAPTTIMPlot.png",sep=""))
  plot(c(1, 200), c(1, 1250), type = "n", main= paste("Subject ",subject.vec[index],"
CAPTTIM",sep=""),
      xlab="Trial",ylab="Regret Per Trial") #Creat a blank plot
  color.index <- data.table:::uniqlist(list(test$Color))
  i <- 1
  while(i < max(color.index)){
    #browser()
    #cat("i is now",i)
    tmp <- which(color.index==i)
    if(length(tmp)==0){
      i <- i+1
      tmp <- which(color.index==i)
    }
    if(length(tmp)==1){
      if(i < max(color.index)){
        if(color.index[tmp+1]-color.index[tmp]==1){ #check for single change points
at a trial
          #cat("i is",i,"\n")
          rect(color.index[tmp],0,color.index[tmp+1],100,col=test$Color[i])
          i <- i+1
          tmp <- which(color.index==i)
        }
      }
      if(length(tmp)!=0 && tmp !=length(color.index)){
        if(color.index[tmp+1]-color.index[tmp]==1){ #check for single change points
at a trial
          #cat("i is",i,"\n")
          rect(color.index[tmp],0,color.index[tmp+1],100,col=test$Color[i])
          i <- i+1
          next
        }
        if(color.index[tmp+1]-color.index[tmp]>1){
          #cat("i is",i,"\n")
          rect(color.index[tmp],0,color.index[tmp+1],100,col=test$Color[i])
          i <- i+1
          tmp <- which(color.index==i)
```

```
      }
    }
    if(length(tmp)!=0 && tmp == length(color.index)){
      rect(color.index[tmp],0,200,100,col=test$Color[i])
      break
    }
    else{
      #cat("i is",i,"\n")
      i <- i+1
    }
  }

 }

 lines(test$regret.trial,lty=2,col="blue")
 dev.off()


}

Regret.mb.df$Regret.Level <- subject.control.vec1
Regret.mb.df$Capttim.Category <- subject.category1
save.image("C:/Users/John/Documents/NPS/Thesis/ThesisData/Data
Critz/RegretData.RData")
write.csv(Regret.mb.df,file="SubjectData.csv")
```

D. CORRELATION TEST R SCRIPT

```
#Loop through each subject
#Take out row 16 of MA summary
MA.summaryTest <- MA.summary[-16,]
red.count.vec <- vector()
green.count.vec <- vector()
for(i in MA.summaryTest$Subject){
 tmp.df <- Regret.mb.df[Regret.mb.df$subject==i,]
 red.count <- sum(tmp.df$Capttim.Category=='red')
 red.count.vec <- c(red.count.vec,red.count)
 green.count <- sum(tmp.df$Capttim.Category=='green')
 green.count.vec <- c(green.count.vec, green.count)
}

pearsonTest(red.count.vec,MA.summaryTest$mb.FD.200)

pearsonTest(red.count.vec,MA.summaryTest$mb.adv.sb.200)

spearmanTest(green.count.vec,MA.summaryTest$mb.FD.200)
```

```
spearmanTest(green.count.vec,MA.summaryTest$mb.adv.sb.200)

png(paste("CorrelationTestRedFD.png"))
plot(xlab = "Number of Trials in Red CAPTTIM Category",
    ylab = "Final Damage Score",
    red.count.vec,
    MA.summaryTest$mb.FD.200, col = "red")
dev.off()

png(paste("CorrelationTestRedAdvSelectBias.png"))
plot(xlab = "Number of Trials in Red CAPTTIM Category",
    ylab = "Advantageous Selection Bias",
    red.count.vec,
    MA.summaryTest$mb.adv.sb.200, col = "red")
dev.off()

png(paste("CorrelationTestGreenFD.png"))
plot(xlab = "Number of Trials in Green CAPTTIM Category",
    ylab = "Final Damage Score",
    green.count.vec,
    MA.summaryTest$mb.FD.200, col = "green")
dev.off()

png(paste("CorrelationTestGreenAdvSelectBias.png"))
plot(xlab = "Number of Trials in Green CAPTTIM Category",
    ylab = "Advantageous Selection Bias",
    green.count.vec,
    MA.summaryTest$mb.adv.sb.200, col = "green")
dev.off()
```

E. EXECUTE R SCRIPT

```
#Workflow
rm(list=ls())
setwd("~/NPS/Thesis/Thesis Data/Data Critz")
source('ReviesedMultiArm_Scrub.v13_Critz.R')
source('RegretMeanPlots_Critz.R')
source('RectangleFinalPlot_Critz.R')
save.image('FinalDataScrub.RData')
```

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

Agrawal, R. (1995). Sample mean based index policies with O(log n) regret for the multi-armed bandit problem. *Advances in Applied Probability, 27*(4), 1054–1078, 1995.

Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, *50*(1), 7–15.

Bell, D.E. (1982) Regret in decision-making under uncertainty. *Operations Research 30*(5):961–981. http://dx.doi.org/10.1287/opre.30.5.961

Braimah, O. J., Osanaiye P. A., Omaku P. E., Saheed, Y. K., & Eshimokhai, S.A. (2014). On the use of exponentially weighted moving average (Ewma) control chart in monitoring road traffic crashes. *International Journal of Mathematics and Statistics Invention (IJMSI), 2*(5), 01–09.

Crowder, S. V. (1989). Design of exponentially weighted moving average schemes. *Technometrics, 21*(3), 155162.

Djulbegovic, B., Elqayam, S., Reljic, T., Hozo, I., Miladinovic, B., Tsalatsanis, A., Kumar, A., Beckstead, J., Taylor, S., & Cannon-Bowers, J. (2014). How do physicians decide to treat: an empirical evaluation of the threshold model. *BMC Medical Informatics and Decision-making* 2014, 14, 47.

Fricker, R. (2010). *Introduction to Statistical Methods for Biosurveillance* (1st Ed.). New York, NY: Cambridge University Press.

Kalgonda, A. A., Koshti, V. V., & Ashokan, K. V. (2011). Exponentially weighted moving average control chart. *Asian Journal of Management Research, 2*(1), 253–263.

Kennedy, Q., Nesbitt, P., & Alt, J. (2014). Assessment of cognitive components of decision-making with military versions of the IGT and WCST. Human Factors and Ergonomics Society 2014 International Annual Meeting, Oct 27–31, Chicago Illinois.

Killick, R., & Eckley, I. (2014). Changepoint: an r package for change point analysis. *Journal of Statistical Software, 58*(3). Retrieved April 15, 2015, from http://www.jstatsoft.org/v58/i03/paper

Krain, A. L., Wilson, A. M., Arbuckle, R., Castellanos, F. X., & Milham, M. P. (2006). Distinct neural mechanisms of risk and ambiguity: a meta-analysis of decision-making. *NeuroImage*, *32*(1), 477–484. doi:http://dx.doi.org/10.1016/j.neuroimage.2006.02.047

Lopez, T. (2011). Odierno outlines priorities as army chief. *Army News Service.* Retrieved from http://www.defense.gov/News/NewsArticle.aspx?ID=65292

Maidstone, R., Fearnhead, P., & Letchford, A. (2013). Changepoint detection using a segment neighborhood approach. Retrieved May 06, 2015, from http://www.lancaster.ac.uk/pg/maidston/xmasposter2013.pdf

Nesbitt, P., Kennedy, Q., & Alt, J. K. (2015). *Iowa gambling task modified for military domain.* Monterey, CA: Operations Research Department, Naval Postgraduate School.

Nesbitt, P., Kennedy, Q., Alt, J., Yang, J., Fricker, R., Appleget, J., Huston, J., Patton, S., & Whitaker, L. (2013). *Understanding optimal decision-making in wargaming.* Monterey, CA: TRAC-Monterey Technical Report, TRAC-M-TR-13–063, September 1, 2013.

Odierno, R., & McHugh, J. (2015). Army human dimension strategy Draft Version 5.3. US Army Publication.

R Development Core Team (2008). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.Retrieved from http://www.R-project.org

Roberts, S. W. (1959). Control chart tests based on geometric moving averages, *Technometrics,* 1, 239–250.

Sutton, R., & Barto, A. (1998). *Reinforcement learning: An introduction.* Cambridge, MA: MIT Press.

Taylor, W. (2000). Change-point analysis: a powerful new tool for detecting changes. Retrieved April 15, 2015, from http://www.variation.com/cpa/tech/changepoint.html

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California

Appendix D
Analysis of Performance on a Modified Wisconsin Card Sorting Test for the Military

The following pages contain the technical report for the modification and analysis of the WCST by Moten et al. Distribution is unlimited.

Analysis of Performance on a Modified Wisconsin Card Sorting Test for the Military

Cardy Moten III

TRADOC Analysis Center, Monterey

Quinn Kennedy, Jonathan K. Alt, Peter Nesbitt

Operations Research Department, Naval Postgraduate School, Operations Research Department,

Naval Postgraduate School, TRADOC Analysis Center, Monterey

Author Note

This research is funded by the Army Research Office.

Cardy Moten III
Office: (831) 656-2452
Cell: (785) 477-3415
cardy.moten3.mil@mail.mil

700 Dyer Road, Room WA-176
Monterey, CA 93943-0692

Abstract

Background. Current Army doctrine stresses a need for military leaders to have the capability to make flexible and adaptive decisions based on a future unknown environment, location, and enemy. To assess a military decision maker's ability in this context, we modified the Wisconsin Card Sorting Test (WCST), a common psychology measure that assesses cognitive flexibility.

Methods. Thirty-four military officers from all branches of service completed the modified WCST. We scored the participants using conventional WCST scoring measures, and used cluster analysis to assess a participant's cognitive flexibility as a high or low performer and nonparametric Mann-Whitney tests to compare high and low performers on certain scoring measures.

Results. The cluster analysis produced three distinct clusters based on total non-perseverative errors. Nonparametric statistical analysis of a decomposition of non-perseverative error into efficient errors and random errors showed that participants who completed the map task accumulated a lower amount of random errors than participants that did not complete the map task.

Limitations. The study only involved military officers and did not include any enlisted military members.

Conclusion. This study serves as the first step in customizing cognitive psychological tests for a military purpose. Based on our findings, all participants that achieved a shift in sorting rule demonstrated adequate cognitive flexibility. However, participants that did not complete all the required sorting rules changed their sorting strategy too soon within a series, resulting in a high quantity of random errors.

Keywords: Wisconsin Card Sorting Task, military decision making, cognitive flexibility, wargaming

Analysis of Performance on a Modified Wisconsin Card Sorting Test for the Military

The U.S. Army published its operating concept in October of 2014. The purpose of this concept is to describe how the Army will operate at the strategic, operational, and tactical level without knowing much about the future environment, location, and enemy (U.S., Department of the Army Training and Doctrine Command, October 2014). In order to accomplish this objective, the training for Army officers has to focus on adaptive decision making through realistic training in actual and virtual environments (U.S., Department of the Army Training and Doctrine Command, October 2014). The purpose of this study is to develop a task that measures a military relevant conceptual component and evaluates the decision making behavior of active duty military officers on this task. According to Army doctrine, a key conceptual component for Army leader's intellectual ability is mental agility. Mentally agile leaders are able to anticipate and adapt to a given situation in order to make the best decision (U.S., Department of the Army Training and Doctrine Command, August 2012). For example, the type of operations executed in Iraq and Afghanistan required military leaders to daily assess the situation in their environment and make the necessary changes to their tactics for survival (Brown, 2007; Hartman, 2008; Mulbury, 2007). In psychology and neuroscience, this ability is known as cognitive flexibility and has been tested in multiple laboratory environments (Vartanian & Mandel, 2011). Although there are laboratory-based tests that measure cognitive flexibility, they are not directly applicable for military training needs. Gallagher and Prestwich suggest using computer-based games to train military leaders' cognitive adaptability to address this need (2012). As a starting point, we set out a method to modify a current psychological task into a military context in order to gain insight into the decision-making ability of military personnel.

One common psychological task that measures the cognitive flexibility of a decision maker is the Wisconsin Card Sorting Task (WCST) (Grant & Berg, 1948). The WCST taps the working memory, set-shifting, and inhibition components of executive function. Participants view five cards, one card displayed at the top center of each screen, the remaining four displayed across the bottom of the screen. Each card contains symbols that vary in number, shape, and color. Over several trials, participants try to figure out the matching rule that will correctly match the card on top of the screen with one of the four cards at the bottom of the screen. Unbeknownst to participants, the matching rule changes once they have 10 consecutive correct matches. For example, after 10 consecutive correct matches based on the color of the symbols, the matching rule could then change to the number or shape of the symbols. Thus, participants must not only learn and maintain in working memory the correct matching rule while inhibiting irrelevant stimuli, but also exhibit cognitive flexibility in detecting when the rule has changed (Grant & Berg, 1948). The task is completed when either a participant successfully completes two rounds of each matching rule for a total of six rules, or until they completed 128 trials. The main performance measures of the WCST include total percent correct, percent of perseverative responses (the number of incorrect responses that would have been correct for the previous matching rule), percent non-perseverative responses (all other incorrect responses after excluding perseverative errors), the number of matching rules achieved, and total number of trials completed (fewer indicates better performance). Our modified version of the WCST is called the map task, in which participants match maps that contain military graphics.

Therefore, the goal of this study was to analyze the mechanisms of executive function, particularly set-shifting, of military decision makers in a military context. Our participants only receive information by sampling each option and collecting an observation. It is expected that

participants with high levels of cognitive flexibility will achieve more rule changes as they progress through the task. Additionally, cluster analysis allows us to identify distinct groups with selected established performance measures. Identifying high and low performing groups provides a better understanding of what characterizes successful decision making. Also, non-parametric analysis allows us to compare the performance groups on select performance measures.

Method

Participants

The study collected data from a convenience sample, from the Naval Postgraduate School, of 34 military officers from all branches of service: nine U.S. Army, eleven U.S.M.C., ten U.S. Navy, three U.S. Coast Guard, and one U.S. Air Force. Participant mean age was 35.11 years (s = 4.90) and mean time in service was 12.7 years (s = 4.42). The average time deployed was variable (M=19.57, s = 12.12 months) (note: one participant did not report their deployment time). Of the 31 participants with deployment experience, the time since their last deployment was 37.98 months (s = 25.18), and 19 of those deployments were to ground combat zones (Iraq or Afghanistan). Over seventy percent of the participants served as staff officers during their most recent deployment. The majority of the participants were male (30 males, 4 females) and all participants possessed an acceptable level of visual acuity (at least 20/30) to complete the decision task. The participants scored within normal ranges in two sets of cognitive measures of assessing visual processing speed (Tombaugh, 2004; Grant & Berg, 1948) and short term memory (Lezak, 1995; Weschler, 2008). The Trails A mean score of 22.60 seconds (s = 6.29) and the Trails B mean score of 44.04 seconds (s = 20.13) showed normal ranges of performance in visual processing speed. The mean Digit Span Forwards score of 11.44 seconds (s = 2.11) and

mean Digit Span Backwards score of 9.53 seconds (s = 2.43) showed normal ranges of performance in working memory. The race and ethnicity of the participants were not noted. All participants had at least an undergraduate degree.

Decision Task

Map task (Modified WCST). Although the literature shows mixed results on what the WCST is actually testing, some modifications of the WCST were successful in developing a valid executive function measure (Nelson, 1976; Ozonoff, 1995; Barceló & Knight, 2002; Kado, et al., 2012). Furthermore, Miyake et al. (2000) found that set-shifting was the key executive function measurement of the WCST. Based on these findings, our focus for this study will be the executive function of set-shifting which we will term as cognitive flexibility.

On a computer screen, participants saw five maps, in which one map is at the top center of the screen and the remaining four are across the bottom of the screen (see Figure 1). Each map contains military graphics that vary in meaning, color and shape (U.S., Department of the Army, September 2004). Graphics have three different categories distinguishable by their color: friendly force (blue), type of intended action, such as ambush (black), and type of enemy force (red). Each of these categories has three different possible shapes, each shape indicating a particular type of friendly force (rectangle and circle), intended action (lines and arrows), or enemy force (diamond) (see Figure 2). Similar to the method of Nelson (1976), we reduced the matching criteria on the map task to the type of graphic: friendly, intent, or enemy. For example if the current matching rule is friendly graphics and the top card shown is similar to the card in Figure 1, then the correct choice would be to choose the card in the lower left hand corner of Figure 1. One additional modification from the WCST is that not all maps have all three types of graphics and participants can match maps based on the absence of graphic type (see Figure 2).

These modifications should reduce some of the ambiguity associated with the original WCST and provide a military context for test participants.

Participants receive instructions to match one of the four lower maps to the top map over an unknown number of trials. As in the original WCST, instructions are purposely vague. Unbeknownst to the participant, the matching rule changes once the participant has 10 consecutive correct matches. As in the WCST, the task is completed when either the participant has successfully completed two rounds of each matching rule for a total of six rounds or until they have exhausted their available 128 trials.

Decision performance variables are measured using typical WCST variables: Total number of trials (fewer indicate better performance), percent correct, percent perseverative responses (percent of trials in which participants incorrectly used the same matching rule as in their previous selection), percent non-perseverative error, number of trials to complete the first matching rule, and number of rules achieved (max number = 6). Similar to the method of Barceló and Knight (2002), we decompose non-perseverative errors into efficient and random components and tabulate the totals for each participant. Efficient errors are scored when an incorrect response is given during the second trial of a new matching rule series, and random errors are an incorrect response on a trial after the participant achieved a correct response on the previous trial (Barceló & Knight, 2002).

Measures

Demographic survey. Demographic information regarding age, gender, service branch, rank, and deployment experience were captured in the demographic survey.

Post-task survey. The post-task survey included questions regarding the map feature the participants sorted on and how soon they realized a sorting rule change took place. The map

feature question was a free response question, and the sorting rule realization question consisted of the following categories: Immediately/After 1-2 trials; After a few trials (3-4 trials); After several trials (5+ trials); and Did not realize the sorting rule had changed.

Trails A and Trails B. Because the map task places demands on visual processing speed, we included Trails A and Trails B tests as covariate measures of these cognitive functions. Trails A and Trails B test visual processing speed (Weschler, 2008). In Trails A, the numbers 1 through 25 are randomly distributed on a worksheet. The participant starts at 1 and must draw a line to each number in chronological order. Participants are instructed to work as quickly and accurately as they can. In Trails B, participants now see both numbers and letters and must connect 1 to A, A to 2, 2 to B and so on until they reach L then 12. They also are instructed to work as quickly and accurately as they can. The test retest reliability on these measures range from .76 to .94 (Wagner, Helmreich, Dahmen, Klaus, & Tadic, 2011). In the current sample, performance on Trails A and B was moderately correlated, as expected $(r = .506,\ p = .003)$. Trails A and B have age and education based norms; these norms were used in computing Trails A and B performance in the current sample (Tombaugh, 2004).

Digit Span Forwards and Backwards tests. The map task also relies on working memory. Therefore, digit span forwards and backwards test from the Wechsler Adult Intelligence Scale (WAIS-IV) measures working memory (Weschler, 2008). In digit span forwards, the experimenter states a series of digits, starting with two digits, and the participant must repeat them back. The number of digits increases, with two trials per number of digits. The test is discontinued if the participant has an incorrect response to both trials for a particular number of digits. In digit span backwards, the same procedure is followed, except this time the participant must repeat the digits in the reverse order. The maximum number of digits is 16 for forward and

16 for backward by different participants. Test retest reliability of the digit span measures range from .66 to .89 (Lezak, 1995). In the current sample, performance on digit span forwards and backwards has a positive correlation as expected ($r = .350$, $p = .042$).

Visual acuity test. Because the decision tasks are visually based, the Snellen eye chart is used to measure visual acuity at the beginning of the experiment. The Snellen eye chart is placed on the wall and consists of 11 lines of block letters, in which each line of letters gets progressively smaller. Participants stand 20 feet from the chart; cover one eye, read aloud as many lines as they can. They then cover the other eye and read aloud as many lines as they can. The experimenter records the last line that the participant could accurately read for each eye. Participants needed at least 20/30 vision to complete the study.

Environment and Equipment

A purpose built synthetic environment was developed for the study. The participant sat at a standard desk and completed the tasks as if they were informing, yet removed from, tactical operations from a military operations center. The tasks were developed in consultation with military advisors. The tasks were written in Python scripting language and presented on a laptop computer running the Windows 7 operating system.

Statistical Modeling Techniques

Cluster Analysis. Using the result of a factor analysis, we chose to separate the sample of participants into clusters. The k-means algorithm clustered the participants according to the desired measures of performance. The placement of a participant in a group reflects an aggregate and relative assessment of their cognitive ability as a high or low performer.

Mann-Whitney Test. The Mann-Whitney test was used to compare the performance of participants that completed the tasks and participants that did not complete the task on all

assessed measures. All tests were conservatively conducted as two-tailed with a .05 alpha significance level. Effect size was computed by dividing the test statistic by the square root of the tested sample and assessed using Cohen's criteria.

Procedures

The institution's IRB approved the study. Participants attended the laboratory individually for a single testing session. They first completed the approved consent form, then a visual acuity test, demographic survey, Trails A and B and Digit Span tests. Finally, participants completed the map task followed by answering the post-task survey questionnaire.

Results

Table 1 shows the summary statistics of the map task results. We observe that the average value for perseverative response rate and number of rules achieved were consistent with our expectations. The percent correct of responses was below our expectations; the percent non-perseverative error was higher than expected.

Cluster Analysis

The cluster analysis produced three distinct grouping by the performance measure of non-perseverative error using factor analysis and k-means clustering as shown in Figure 3. We chose non-perseverative errors because it was the highest loading variable with a value of 0.99 in the factor analysis. The first cluster consists of 14 participants with a high number of non-perseverative errors; the second cluster contains 12 participants with a moderate number of non-perseverative errors; and the third cluster has eight participants with a low number of non-perseverative errors. None of the participants in the first cluster, one participant (ID #7) in the second cluster, and all participants in the third cluster completed all six-rule changes of the map task.

Based on these results, we further classify participants as high or low performers. High performers are participants that completed all six rule changes, and low performers are participants that did not complete all six rule changes. The nine high performers had a total of non-perseverative errors $(M = 14.11)$ that were significantly lower than the total of non-perseverative errors for low performers $(M = 51.84,\ z = -4.4,\ p < .0001,\ \text{effect size} = 0.753)$. The number of trials to complete a rule can indicate a single participant's or groups of participants' ability to complete the map task within the allotted amount of trials. The total number of trials for high performers $(M = 95.33)$ was significantly lower than the total amount of trials for low performers $(M = 128,\ \text{z} = -5.6,\ \text{p<.0001, effect size} = 0.970)$. Figure 4 shows the number of trials participants required to complete the first matching rule clustered by total non-perseverative errors. Only 10 of the 14 participants in the first cluster, and all participants in the second and third clusters completed the first matching rule. We found the high performers needed a statistically significant lower amount of trials $(M = 17.89)$ to complete the first matching rule than the lower performers $(M = 53.72,\ \text{z} = -3.4,\ p < .0002,\ \text{effect size} = 0.583)$. Furthermore, Figure 5 indicates that only two of the 14 participants in the first cluster, and all participants in the second and third clusters completed the first three matching rules. Figure 6 displays that only one of the 14 participants in the first cluster, six of the 12 participants in the second cluster, and all eight participants in the third cluster completed the first five matching rules. We conducted correlation test of the non-perseverative errors with the results of the digital span and the Trails A and B test. The results were not significant for the digit span forward test $(r = 0.213,\ p = 0.227)$, digit span backwards test $(r = 0.005,\ p = 0.979)$, Trails A normed percentile $(r = 0.046,\ P = 0.798)$, and Trails B normed percentile $(r = 0.099,\ p = 0.584)$.

Non-perseverative Error Analysis

Table 2 shows the summary statistics for the follow-on analysis of non-perseverative errors. As expected the participants that achieved a change in matching rule did so efficiently. Additionally there was a not a significant difference in the average number of efficient errors for the high performers than for low performers $(z = 1.29, p = 0.21, \text{effect size} = 0.22)$. High performers however did have a significant difference in perseverative errors, $(z = 3.27, p = 0.0006, \text{effect size} = 0.56)$ and random errors $(z = -4.39, p < 0.0001, \text{effect size} = 0.75)$. These results indicate that participants are efficiently finding a new matching rule, but the low performing participants are shifting to a new rule too soon in the current series.

Post-task Survey

We also analyzed the responses of the high and low performers on a question in the post-task survey that asked when they realized that the sorting rule had changed. One high performer reported realizing a rule change within 1-2 trials, and the remaining eight high performers reported realizing a rule change within 3-4 trials. For the low performers, nine reported recognizing a rule change within 1-2 trials; seven within 3-4 trials; four within 5+ trials; and five participants did not realize the rules had changed at all.

Discussion

Military operations require leaders to have agile and adaptive decision making skills. However, current military training typically does not focus on training cognitive functions necessary for optimal decision making. The purpose of this study was to analyze the critical components of set-shifting executive function to potentially fill this current gap in capability. Based on our findings, all participants that achieved a shift in sorting rule demonstrated adequate

cognitive flexibility. However, participants that did not complete all the required sorting rules changed their sorting strategy too soon within a series, resulting in a high quantity of random errors.

Using cluster analysis, we determined that using non-perseverative errors may be a useful assessment tool of cognitive flexibility for Soldiers. Because we can directly observe this metric and cluster participants into distinct groups, we can gauge a participant's performance throughout the task. We found that the participants clustered into three noticeable groups. The first group consisted of participants that had no probability of successfully completing the map task to standard. The second cluster encompassed participants that had a small probability of successfully completing the map task. The final cluster contained participants that had a high probability of successfully completing the map task. Having grouped the participants with a sufficient metric, we wanted to further explore the high rate of non-perseverative errors exhibited in the study participants compared to a normed population. We also found through factor analysis that the high non-perseverative response rates for this particular sample were highly correlated with performance.

Similar to the findings of Barceló and Knight (2002), we found that decomposing non-perseverative errors into efficient and random errors produced further insight into the potential reasons for a participant not completing all six matching rules of the map task. Although all participants that achieved a change in matching rule were consistent in exploring early in a sorting series to determine a matching rule, the lower performing participants, however, ended up accumulating more random errors in the later trials of a sorting series resulting in these participants not completing the map task. Contrary to the original WCST, the military symbols on the map task have a specific meaning and experienced officers could read each card as a

military operation. The symbols on the map task are primarily ground-based, and this could result in officers familiar with these symbols to attempt to match the cards as a type of military operation and not just simply matching on the correct symbol color. If this hypothesis were true, then we might have support to explain the large number of non-perseverative errors, manifested as random errors, measured in the conduct of the map task. The end result of this action is that these officers are switching decision making tactics too soon to see if a particular tactic actually works, and therefore don't reach the point at which they are exploiting on the correct matching rule at that time.

A limitation of this study was that participants were a sample of convenience of military officers from the Naval Postgraduate School who served at least five years in the military. Although the study did have participants from all branches of the Armed Forces, a study that involves not only officers but enlisted military members as well, could provide further insights into the set-shifting of military members of varying years of experience.

A natural extension of this work would be to modify the instructions and reduce the total card pile and determine if the participants would have fewer trials with random errors similar to other studies (Nelson, 1976; Ozonoff, 1995; Barceló & Knight, 2002; Kado, et al., 2012). A regret analysis at a trial-by-trial level could also be conducted to determine the optimal range for a training intervention for a participant that is having difficulty in making the proper set shifts. Additionally, decision-making performance could be analyzed by modifying the map task cards to show a sequence of military actions and the participants are required to arrange the cards in the correct order of events.

Bibliography

Barceló, F., & Knight, R. T. (2002). Both random and perseverative errors underlie WCST deficits in prefrontal patients. *Neuropsychologia, 40*, 349-356.

Brown, R. B. (2007). The Agile-Leader Mind-Set: Leveraging the Power of Modularity in Iraq. *Military Review, 87*(4), 32.

Gallagher, P. S., & Prestwich, S. (2012). Supporting cognitive adaptability through game design. *6th European Conference on Games Based Learning*, 165.

Grant, D. A., & Berg, E. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem. *Journal of experimental psychology, 38*(4), 404.

Hartman, W. J. (2008). *Exploitation Tactics: A Doctrine for the 21st Century.* Fort Leavenworth: US Army Command and General Staff School of Advanced Military Studies.

Kado, Y., Sanada, S., Yanagihara, M., Ogino, T., Ohno, S., Watanabe, K., . . . Ohtsuka, Y. (2012). Executive function in children with pervasive developmental disorder and attention-deficit/hyperactivity disorder assessed by the Keio version of the Wisconsin card sorting test. *Brain and Development, 34*(5), 354-359.

Lezak, M. D. (1995). *Neuropsychological Assessment.* Oxford University Press.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The Unity and Diversity of Executive Functions and Their Contributions to Complex "Frontal Lobe" Tasks: A Latent Variable Analysis. *Cognitive Psychology, 41*(1), 49-100.

Mulbury, D. S. (2007). *Developing Adaptive Leaders, A Cultural Imperative.* Carlisle Barracks: US Army War College.

Nelson, H. E. (1976). A modified card sorting test sensitive to frontal lobe defects. *Cortex, 12*(4), 313-324.

Ozonoff, S. (1995). Reliability and validity of the Wisconsin Card Sorting Test in studies of autism. *Neuropsychology, 9*(4), 491.

Tombaugh, T. N. (2004). Trail Making Test A and B: normative data stratified by age and education. *Archives of Clinical Neuropsychology, 19*(2), 203-214.

U.S., Department of the Army. (September 2004). *FM 1-02, Operational Terms and Graphics.* Washington D.C.: Government Printing Office.

U.S., Department of the Army Training and Doctrine Command. (August 2012). *ADRP 6-22, Army Leadership.* Washington D.C.: Government Printing Office.

U.S., Department of the Army Training and Doctrine Command. (October 2014). *TRADOC Pamphlet 525-3-1, The U.S. Army Operating Concept: Win in a Complex World.* Washington D.C.: Government Printing Office.

Vartanian, O., & Mandel, D. R. (2011). *Neuroscience of decision making.* Psychology Press.

Wagner, S., Helmreich, I., Dahmen, N., Klaus, L., & Tadic, A. (2011). Reliability of three alternate forms of the trail making tests a and B. *Archives of Clinical Neuropsychology, 26*(4), 314-321.

Weschler, D. (2008). Wechsler adult intelligence scale--Fourth Edition (WAIS--IV). San Antonio, TX: NCS Pearson.

Tables

Table 1

Summary statistics of map task decision performance variables.

	Mean	SD
Number of trials completed	119.35	16.53
Percent correct	0.65	0.15
Percent perseverative errors	0.06	0.08
Percent non-perseverative errors	0.34	0.16
Number of trials to complete first matching rule	42.97	28.95
Number of rules achieved	3.20	1.93

Table 2

Mean and standard deviation of error trials per map task cluster and performance group.

	Efficient		Perseverative		Random	
	Mean	SD	Mean	SD	Mean	SD
Cluster 1 (n = 14)	1.36	1.01	0.71	0.91	62.57	11.73
Cluster 2 (n = 12)	2.33	1.43	2.92	1.08	31.75	5.12
Cluster 3 (n=8)	2.25	1.16	3.75	0.71	9.00	4.17
High Performers (n = 9)	2.33	1.12	3.67	0.71	10.67	6.34
Low Performers (n = 25)	1.76	1.30	1.68	1.49	49.32	17.79

Figures

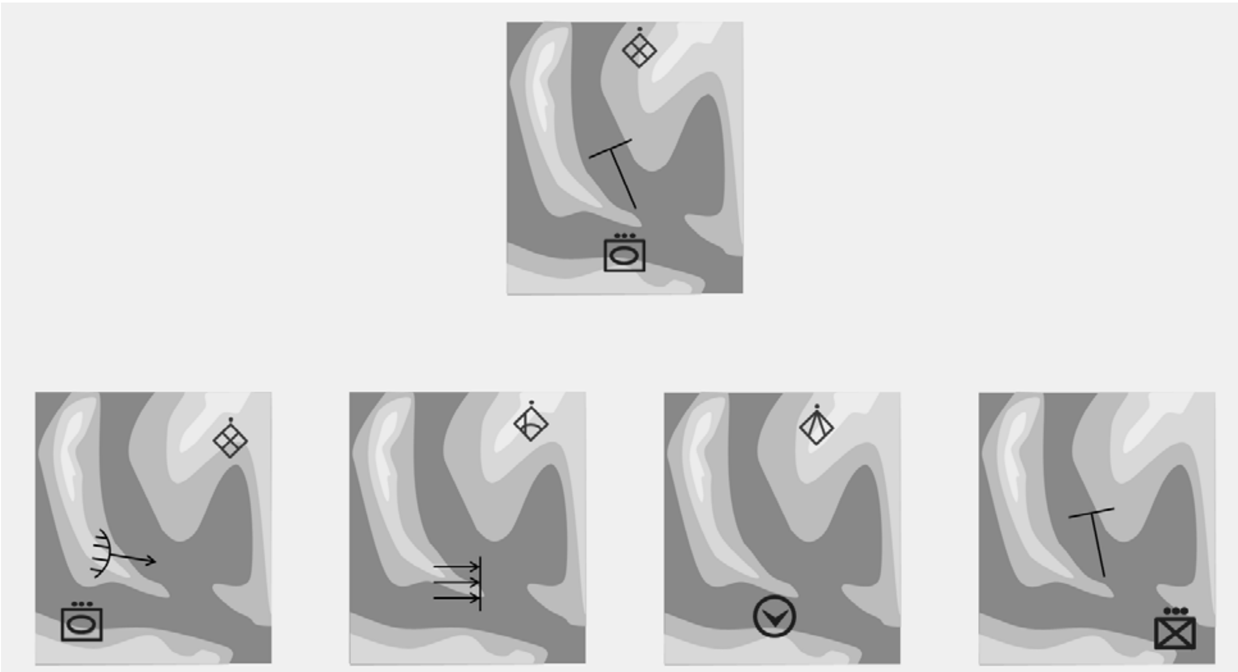


Figure 1. Screen shot of a typical participant's view of the map task. In this example, if the current rule is to select on the enemy (diamond shape) symbol on the top card, then the correct choice is the first card in the bottom row.

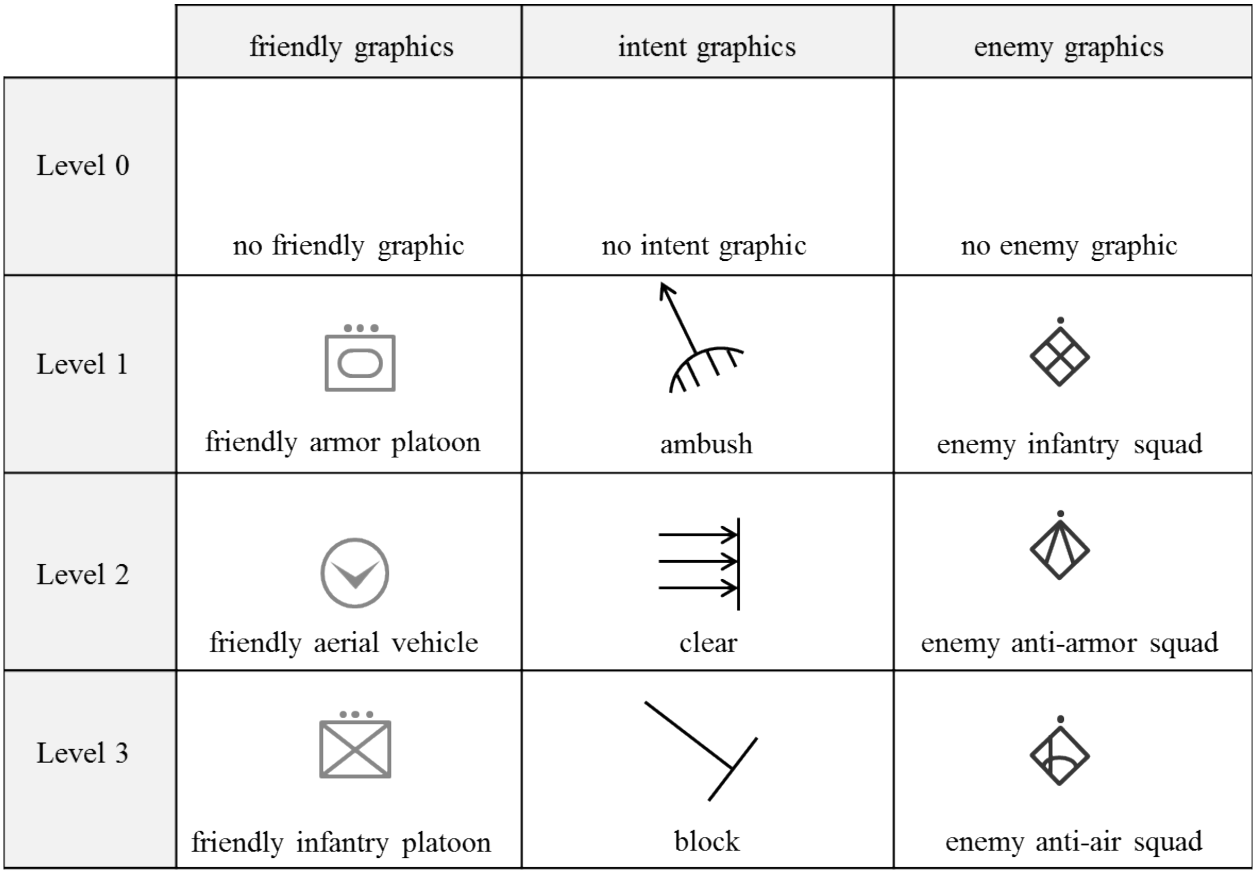


Figure 2. A table of icons used in the map task. We modify the WCST cards to include a matching rule based on the absence of graphics (Level 0).

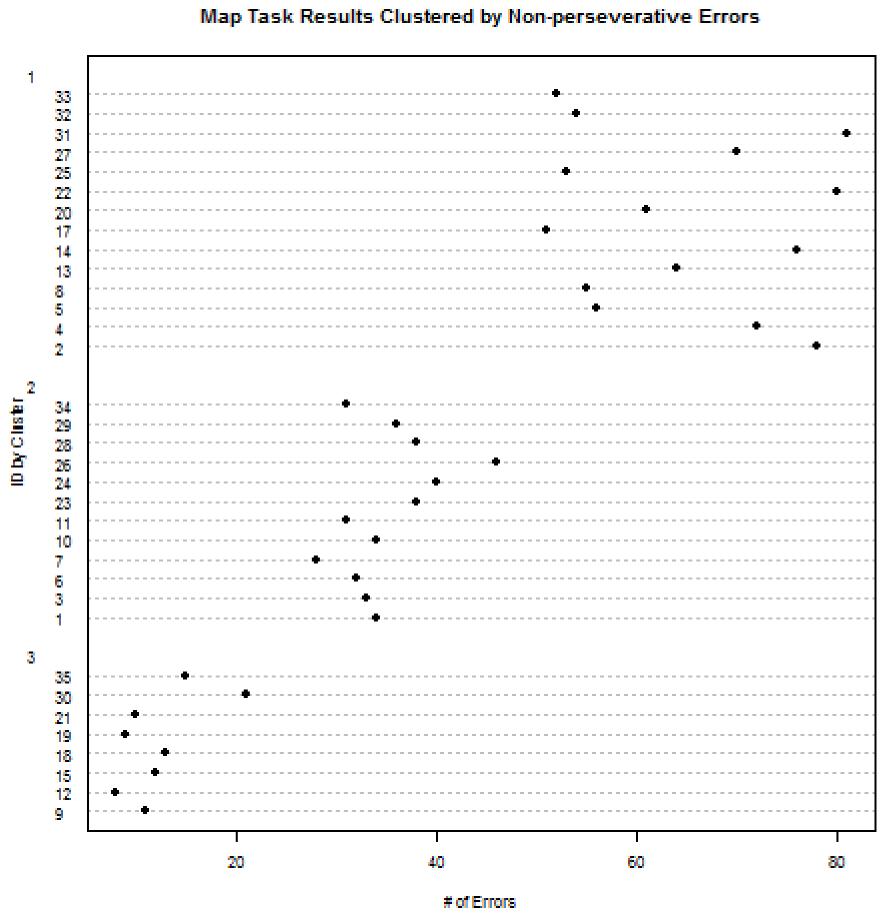


Figure 3. A dot plot map task participant clusters. In this figure, we can see three distinct groups of participants. Cluster one, two, and three represent participants with a high, moderate, and low number of non-perseverative errors respectively.

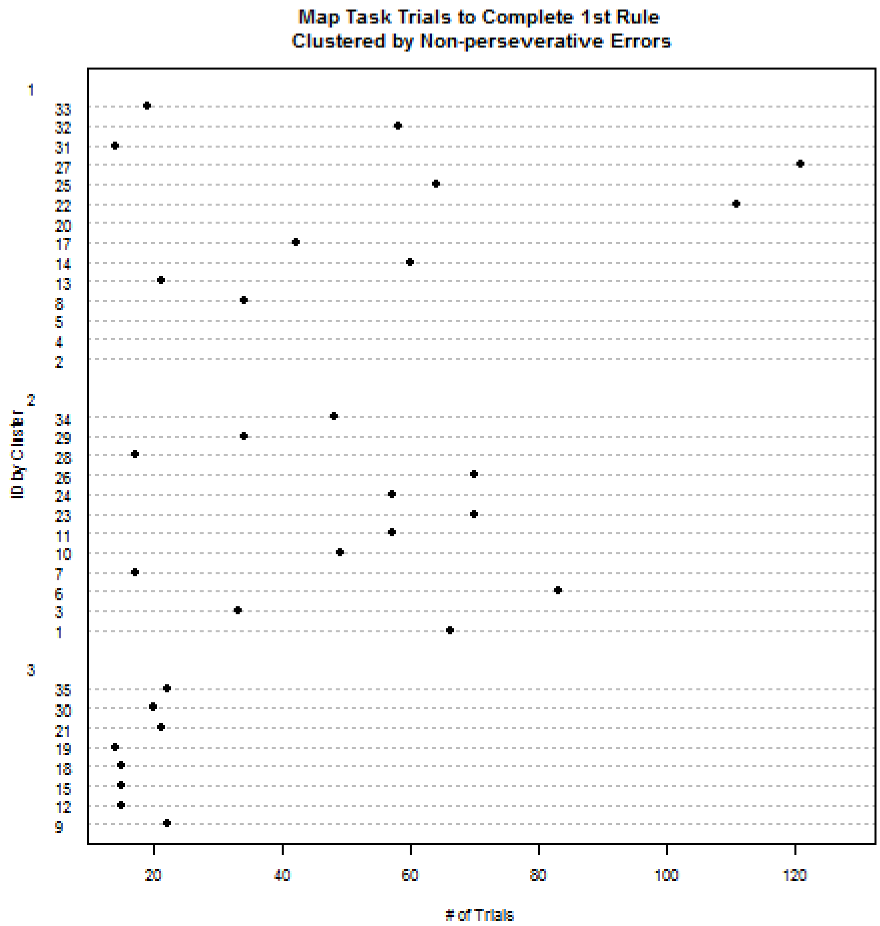


Figure 4. A dot plot of the number of trials a participant required to complete the first matching rule. The participants are grouped by the same clusters as shown in Figure 3. We notice that 10 of 14 participants in cluster one, all 12 participants in cluster two, and all eight participants in cluster three completed the first matching rule.

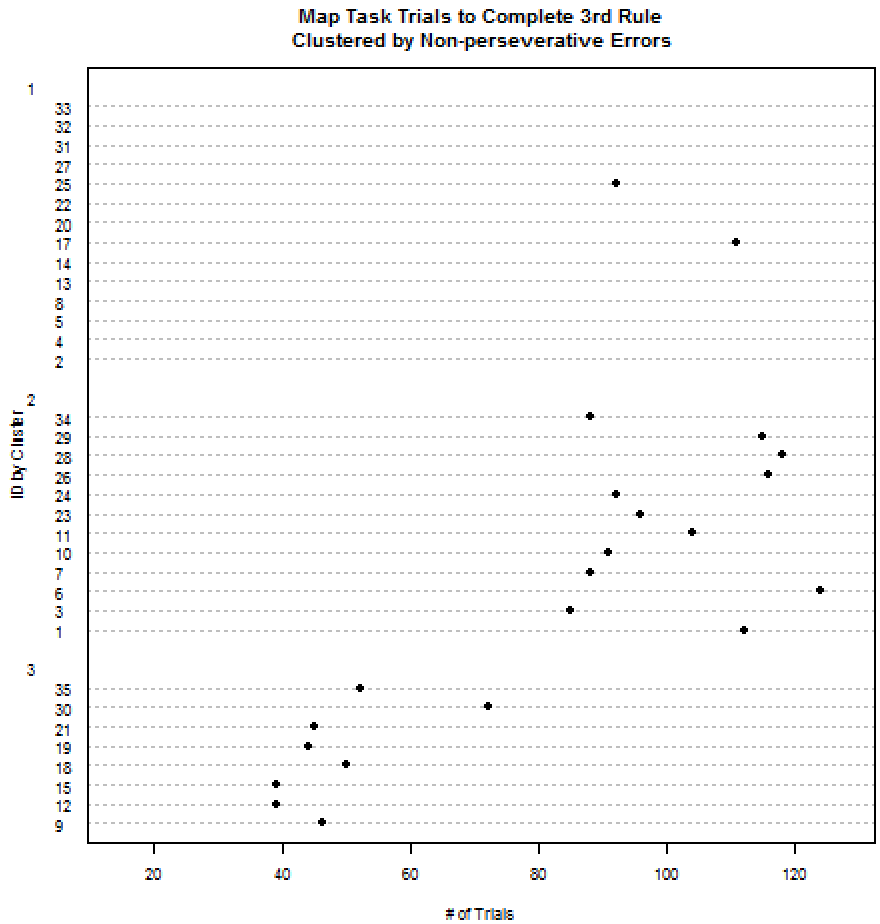


Figure 5. A dot plot of the number of trials a participant required to complete the third matching rule. The participants are grouped by the same clusters as shown in Figure 3. We notice that 2 of 14 participants in cluster one, all 12 participants in cluster two, and all eight participants in cluster three completed the first matching rule.

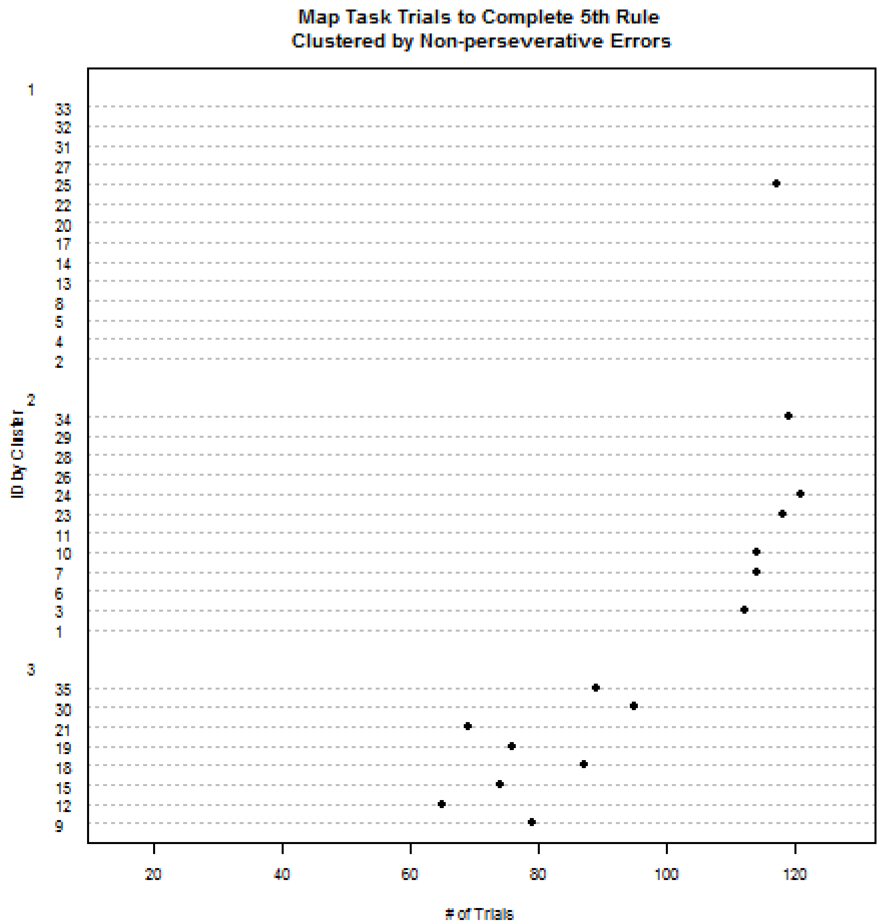


Figure 6. A dot plot of the number of trials a participant required to complete the fifth matching rule. The participants are grouped by the same clusters as shown in Figure 3. We notice that one of 14 participants in cluster one, 6 of the 12 participants in cluster two, and all eight participants in cluster three completed the first matching rule.

Appendix E
References

[1] U.S., Department of the Army Training and Doctrine Command. *TRADOC Pamphlet 525-3-1, The U.S. Army Operating Concept: Win In a Complex World*. Washington DC: Government Printing Office, October 2014.

[2] Antoine Bechara et al. "Insensitivity to future consequences following damage to human prefrontal cortex". In: *Cognition* 50.1 (1994), pp. 7–15.

[3] David A Grant and Esta Berg. "A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem." In: *Journal of experimental psychology* 38.4 (1948), p. 404.

[4] Quinn Kennedy, Peter Nesbitt, and Jon Alt. "Assessment of Cognitive Components of Decision Making with Military Versions of the IGT and WCST". In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Vol. 58. 1. SAGE Publications. 2014, pp. 300–304.

[5] Quinn Kennedy et al. *Cognitive Alignment with Performance Targeted Training Intervention Model: CAPTTIM*. Tech. rep. Monterey, California. Naval Postgraduate School, 2015.

[6] Peter Nesbitt et al. "Iowa Gambling Task modified for military domain." In: *Military Psychology* 27.4 (2015), p. 252.

[7] Quinn Kennedy et al. "Can a simple wargame provide an unobtrusive indicator of TBI?" In: *43rd Annual Meeting of the International Neuropsychological Society (INS)*. 2015.

[8] John Critz. "Understanding Optimal Decision Making". MA thesis. Naval Postgraduate School, 2015.

[9] Cardy Moten III et al. "Analysis of Performance on a Modified Wisconsin Card Sorting Test for the Military".

[10] Quinn Kennedy et al. "Cognitive alignment with performance targeted training intervention model: CAPTTIM". In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. 2015.

This page intentionally left blank.

Appendix F
Glossary

ARO	Army Research Office
CAPTTIM	Cognitive Alighnment With Performance Targeted Training Intervention Model
IGT	Iowa Gambling Task
TRAC	Training and Doctrine Command Analysis Center
WCST	Wisconsin Card Sorting Test